UNIVERSITY OF HELSINKI REPORT SERIES IN PHYSICS

HU-P-D170

Identification of Hadronically Decaying Tau Leptons in Searches for Heavy MSSM Higgs Bosons with the CMS Detector at the CERN LHC

Lauri A. Wendland

Division of Elementary Particle Physics
Department of Physics
Faculty of Science
University of Helsinki
and
Helsinki Institute of Physics
Helsinki, Finland

Academic dissertation

To be presented, with the permission of the Faculty of Science
of the University of Helsinki, for public criticism
in the auditorium E207 of the Physicum building, Gustaf Hällströmin katu 2,
on Friday, November 27$^{\text{th}}$, 2009, at 10 o'clock a.m.

Helsinki 2009

Cover image:
The process $\mathrm{gg} \rightarrow \bar{\mathrm{t}}\mathrm{bH}^{+}(312\ \mathrm{GeV}/c^{2})$, $\bar{\mathrm{t}} \rightarrow \mathrm{W}^{-}\bar{\mathrm{b}} \rightarrow \mathrm{s}\bar{\mathrm{c}}\bar{\mathrm{b}}$, $\mathrm{H}^{+} \rightarrow \tau^{+}\nu_{\tau} \rightarrow \mathrm{hadr.}$ (with a radiative gluon jet) simulated inside the CMS detector and visualized in the r, ϕ projection. The curved lines correspond to the reconstructed tracks of electrically charged particles and the rectangular bars visualize the amount of deposited energy in the electromagnetic and hadronic calorimeters.
The image has been generated by the author.

ISBN 978-952-10-5644-4 (printed version)
ISSN 0356-0961
Helsinki University Printing House (Yliopistopaino)

ISBN 978-952-10-5645-1 (PDF version)
http://ethesis.helsinki.fi
Electronic Publications @ University of Helsinki (Helsingin yliopiston verkkojulkaisut)

Helsinki 2009

Acknowledgements

This doctoral thesis is based on research done during the years 2006-2009 at the Division of Elementary Particle Physics of the Department of Physics at the University of Helsinki, at the CMS Physics Analysis group of the Helsinki Institute of Physics (HIP), and at the CMS experiment at the European Center for Particle Physics (CERN). The work has been funded by the Helsinki Institute of Physics, the University of Helsinki, and the Vilho, Yrjö, and Kalle Väisälä fund of the Finnish Academy of Science and Letters. Travel to CERN and to conferences has been funded by the Magnus Ehrnrooth Foundation and the Waldemar von Frenckell Foundation.

During the years I have had the privilege of receiving support and collaborating with many physicists, both at Helsinki and at CERN. I would like to thank my supervisors, Prof. Heimo Saarikko and Prof. Jorma Tuominiemi, for their support and for taking me in as a young physics student in the first place, Doc. Veikko Karimäki for his support and for giving me the opportunity of doing research in the HIP CMS Physics Analysis group, and my colleagues in the HIP CMS Programme and in the CMS Higgs and tau working groups for their help and support. I would especially like to thank Doc. Ritva Kinnunen, Dr. Alexandre Nikitenko, and Dr. Sami Lehti for always having time for questions and for providing valuable comments and support. I would also like to thank Doc. Kenneth Österberg and Dr. Guillaume Unal for reviewing the manuscript and for providing feedback.

I would like to thank all my friends at the OPKO student mission and at the Finnish Luther Foundation for, amongst other things, helping me to keep a healthy balance between work and social life. Last, but not least, I would like to thank my family, especially my parents Marketta and Lothar, for a warm and encouraging home and my dear wife Tiina for always encouraging me to complete this work with her love, willingness to listen, and patience.

From a philosophical point of view it is interesting to be conducting research on a hypothetical particle that, despite reasonable theoretical arguments, may or may not exist. As a colleague put it, spending half of a life time looking for such a particle is fascinating, but it does not give a meaning to life, nor does it provide much comfort if loved ones around you die.

As experimental physicists, we are looking for a fundamental truth that can be found in nature. If the Higgs boson is found, it will be one of the greatest discoveries in modern science. Yet, for me, the greatest discovery during my years at the university has not been the fact that the truth is out there to be found, but that He who claims to be the truth, the way, and the life has actually found me. And this is what gives me the joy, the relevance, and the meaning to marvel and look for truth in nature, as well as the hope of life after death.

"When I look at the sky, which you have made, at the moon and the stars, which you set in their places - what are human beings, that you think of them; mere mortals, that you care for them?" - Psalm 8, the Bible

Soli Deo Gloria - To God alone be the glory.

L. Wendland: Identification of Hadronically Decaying Tau Leptons in Searches for Heavy MSSM Higgs Bosons with the CMS Detector at the CERN LHC, University of Helsinki, 2009, 205 p.+appendices, University of Helsinki, Report Series in Physics, HU-P-D170, ISSN 0356-0961, ISBN 978-952-10-5644-4 (printed version), ISBN 978-952-10-5645-1 (PDF version).

Classification (INSPEC): A1110N, A1130P, A1130Q, A1335, A1338, A1385N, C6185

Keywords (INSPEC): supersymmetry, spontaneous electroweak symmetry breaking, very high-energy p-p interactions, decays of Higgs bosons, decays of tau leptons

Abstract

This thesis describes methods for the reliable identification of hadronically decaying τ leptons in the search for heavy Higgs bosons of the minimal supersymmetric standard model of particle physics (MSSM). The identification of the hadronic τ lepton decays, i.e. τ-jets, is applied to the $\mathrm{gg} \to \mathrm{b\bar{b}H(A)}$, $\mathrm{H, A} \to \tau\tau$ and $\mathrm{gg} \to \mathrm{tbH^{\pm}}$, $\mathrm{H^{\pm}} \to \tau^{\pm}\nu_{\tau}$ processes to be searched for in the CMS experiment at the CERN Large Hadron Collider. Of all the event selections applied in these final states, the τ-jet identification is the single most important event selection criterion to separate the tiny Higgs boson signal from a large number of background events.

The τ-jet identification is studied with methods based on a signature of a low charged track multiplicity, the containment of the decay products within a narrow cone, an isolated electromagnetic energy deposition, a non-zero τ lepton flight path, the absence of electrons, muons, and neutral hadrons in the decay signature, and a relatively small τ lepton mass compared to the mass of most hadrons. Furthermore, in the $\mathrm{H^{\pm}} \to \tau^{\pm}\nu_{\tau}$ channel, helicity correlations are exploited to separate the signal τ jets from those originating from the $\mathrm{W^{\pm}} \to \tau^{\pm}\nu_{\tau}$ decays. Since many of these identification methods rely on the recons truction of charged particle tracks, the systematic uncertainties resulting from the mechanical tolerances of the tracking sensor positions are estimated with care.

The τ-jet identification and other standard selection methods are applied to the search for the heavy neutral and charged Higgs bosons in the $\mathrm{H, A} \to \tau\tau$ and $\mathrm{H^{\pm}} \to \tau^{\pm}\nu_{\tau}$ decay channels. For the $\mathrm{H^{\pm}} \to \tau^{\pm}\nu_{\tau}$ channel, the τ-jet identification is redone and optimized with a recent and more detailed event simulation than previously in the CMS experiment. Both decay channels are found to be very promising for the discovery of the heavy MSSM Higgs bosons.

The Higgs boson(s), whose existence has not yet been experimentally verified, are a part of the standard model and its most popular extensions. They are a manifestation of a mechanism which breaks the electroweak symmetry and generates masses for particles. Since the $\mathrm{H, A} \to \tau\tau$ and $\mathrm{H^{\pm}} \to \tau^{\pm}\nu_{\tau}$ decay channels are important for the discovery of the Higgs bosons in a large region of the permitted parameter space, the analysis described in this thesis serves as a probe for finding out properties of the microcosm of particles and their interactions in the energy scales beyond the standard model of particle physics.

Abstract in Finnish

Tämä väitöskirja käsittelee menetelmiä hadronisesti hajoavien τ-leptoneiden luotettavaan tunnistamiseen raskaiden Higgsin bosonien etsinnässä minimaalisessa supersymmetrisessä standardimallissa (MSSM). Hadronisten τ-leptonien hajoamisten, eli τ-ryöppyjen tunnistamista sovelletaan $\mathrm{gg} \to \mathrm{b\bar{b}H(A)}$, $\mathrm{H, A} \to \tau\tau$ ja $\mathrm{gg} \to \mathrm{tbH}^{\pm}$, $\mathrm{H}^{\pm} \to \tau^{\pm}\nu_{\tau}$ prosesseihin, joita tullaan etsimään CERN:n LHC-kiihdyttimen CMS-kokeessa. τ-ryöppyjen tunnistaminen on tärkein yksittäinen analyysimenetelmä, jolla voidaan erottaa pieni Higgsin bosonin signaali suuresta määrästä taustatapahtumia.

Väitöksessä tutkitaan τ-ryöppyjen tunnistamismenetelmiä, jotka perustuvat pieneen määrään sähköisesti varattuja hiukkasia τ-leptonin hajoamistuotteiden joukossa, hajoamistuotteiden rajautumiseen kapeaan kartioon, eristettyyn sähkömagneettiseen energiaprofiiliin, τ-leptonin lentomatkan pituuteen, τ-leptonin suhteellisen pieneen massaan verrattuna useimpien hadronien massaan sekä elektronien, myonien ja neutraalien hadronien puuttumiseen τ-leptonin hajoamistuotteista. Lisäksi $\mathrm{H}^{\pm} \to \tau^{\pm}\nu_{\tau}$ -kanavassa käytetään hyväksi τ-leptonien helisiteettikorrelaatiota erottamaan signaalin τ-ryöpyt $\mathrm{W}^{\pm} \to \tau^{\pm}\nu_{\tau}$ -hajoamisissa syntyvistä τ-ryöpyistä. Koska monet näistä tunnistusmenetelmistä perustuvat varattujen hiukkasten jälkien rekonstruoimiseen, jälki-ilmaisimen sensorien mekaanisten siirtymien vaikutusta systemaattisiin epävarmuuksiin arvioidaan huolella.

τ-ryöppyjen tunnistamista ja muita yleisesti käytettyjä kokeellisia valintamenetelmiä sovelletaan neutraalien sekä varattujen raskaiden Higgsin bosonien etsimiseen $\mathrm{H, A} \to \tau\tau$ ja $\mathrm{H}^{\pm} \to \tau^{\pm}\nu_{\tau}$ -hajoamiskanavissa. $\mathrm{H}^{\pm} \to \tau^{\pm}\nu_{\tau}$ -kanavassa τ-ryöppyjen tunnistaminen analysoidaan ja optimoidaan uudelleen hiljattain tehdyllä aiempaa tarkemmalla simulaatiolla, Molempien hajoamiskanavien havaitaan olevan hyvin lupaavia MSSM-teorian raskaiden Higgsin bosonien löytämiseksi.

Higgsin bosoni(t), joiden olemassaoloa ei toistaiseksi ole kokeellisesti vahvistettu, ovat osa hiukkasfysiikan standardimallia ja sen lupaavimpia laajennuksia. Ne ovat osa mekanismia, joka rikkoo sähköheikon vuorovaikutuksen symmetrian ja tuottaa massan hiukkasille. Koska $\mathrm{H, A} \to \tau\tau$ ja $\mathrm{H}^{\pm} \to \tau^{\pm}\nu_{\tau}$ -hajoamiskanavat ovat tärkeitä Higgsin bosonien löytämiseksi suuressa osassa sallittua parametriavaruutta, tässä väitöskirjassa esitetty analyysi mahdollistaa hiukkasten ja niiden välisten vuorovaikutusten muodostaman mikrokosmoksen ominaisuuksien kokeellisen tutkimisen energiaskaalalla, joka ylittää hiukkasfysiikan standardimallin tähänastiset pätevyysrajat.

Publications

1. S. Gennai, F. Moortgat, L. Wendland, A. Nikitenko, S. Wakefield, G. Bagliesi, S. Dutta, A. Kalinowski, M. Konecki and D. Kotlinski:
   "Tau jet reconstruction and tagging with CMS",
   Eur.Phys.J. C46 s01 (2006), pp. 1-21

2. CMS Collaboration (incl. L. Wendland):
   "CMS Physics Technical Design Report, Volume I: Detector Performance and software",
   CERN/LHCC 2006-001 (2006), pp. 1-548

3. S. Gennai, A. Nikitenko and L. Wendland:
   "Search for MSSM Heavy Neutral Higgs Boson in tau tau -> two Jet Decay Mode",
   CMS Note 2006/126 (2006), pp. 1-21

4. CMS Collaboration (incl. L. Wendland):
   "CMS Physics Technical Design Report, Volume II: Physics Performance",
   J.Phys. G: Nucl.Part.Phys. 34 (2007), pp. 995-1579

Author's contribution

The research was carried out at the Helsinki Institute of Physics (HIP) and within the Particle Flow & τ Identification and Higgs groups of the CMS collaboration during the years 2006-2009. The author participated actively in the physics analysis of the large Monte Carlo data samples and the development of the analysis code.

Publications 1 and 3 were essentially included in the physics technical design reports (Publications 2 and 4, respectively). The author's contributions to Publications 1 and 2 were the τ tagging methods based on the impact parameter and the τ flight path including secondary vertex reconstruction. In Publications 3 and 4, the author contributed the evaluation of systematic uncertainties to the τ tagging resulting from misalignment effects.

In addition to these publications, the author is the main author of two papers which have been approved by the CMS collaboration as analysis notes which are intended for internal use. The contents of these papers are included in Sections 2.3.2, 5.1.11-5.1.13.1, and 7.7.1. The CMS collaboration approved the inclusion of the results of these papers in this thesis in a plenary meeting on the 29^{th} of June 2009.

Table of Contents

List of acronyms and symbols

APE	Alignment position error
AVF	Adaptive vertex fitter
BR	Branching ratio
CERN	European laboratory for particle phycics
CKF	Combinatorial Kalman filter (used in track finding)
CMS	Compact muon solenoid
CMSSW	CMS reconstruction and analysis software
CMSIM	CMS simulation package
COBRA	Coherent object-oriented base for reconstruction, analysis and simulation
CSC	Cathode strip chamber (for muon measurement)
DT	Drift tube (for muon measurement)
ECAL	Electromagnetic calorimeter
HCAL	Hadronic calorimeter
HEP	High-energy physics
HLT	High-level trigger
IP	Impact parameter
$\mathrm{IP_T}$	Impact parameter in the transverse $\mathrm{x,y}$ plane
KF	Kalman filter (least squares method to reconstruct tracks)
KVF	Kalman vertex fitter
L1	Level-1 trigger
LEP	Large electron-positron collider
LHC	Large hadron collider
LO	Leading order (i.e. tree-level) perturbative terms in terms of coupling constant
MC	Monte-Carlo (simulation)
MSSM	Minimal supersymmetric standard model (of particle physics)
NLO	Next to leading order perturbative terms in terms of coupling constant
ORCA	Object-oriented reconstruction (software) for CMS analysis
OSCAR	Object oriented simulation for CMS analysis and reconstruction
PS	Proton synchrotron
PU	Pile-up (events)
PV	Primary vertex
QCD	Quantum chromodynamics
QED	Quantum electrodynamics
RHIC	Relativistic heavy ion collider
RPC	Resistive plate chamber (for muon measurement)
SM	Standard model (of particle physics)
SPS	Super proton sychrotron
SSB	Spontaneous symmetry breaking
SV	Secondary vertex
SUSY	Supersymmetry model of particle physics
TDR	Technical design report
TKF	Trimmed Kalman fitter (for vertex fitting)

3D	Three-dimensional
$\mathrm{E_T}$	Energy in the transverse x,y plane
$\mathrm{h.c.}$	Hermitian conjugate
$\mathcal{L}$	Lagrangian (density) or luminosity
p	Momentum
$\mathrm{p_T}$	Momentum in the transverse x,y plane
r	Coordinate transverse to the beam axis, $\mathrm{r}^2 = \mathrm{x}^2 + \mathrm{y}^2$
$\Delta\mathrm{R}$	Distance in the η,ϕ plane, $\Delta\mathrm{R} = \sqrt{\Delta\eta^2 + \Delta\phi^2}$
x	Horizontal coordinate transverse to the beam axis
y	Vertical coordinate transverse to the beam axis
z	Coordinate parallel to the beam axis
η	Pseudo-rapidity ($\eta = -\ln\tan\frac{\theta}{2}$)
ϕ	Azimuthal angle in the x,y plane between the x and y components
σ	Standard deviation
θ	Angle in the y,z plane between the y and z components
τ jet	Hadronically decaying τ lepton

1 Introduction

The goal of elementary particle physics is to study the smallest constituents of matter and how they interact with each other. To study this fascinating microcosm, highly complex devices with large dimensions have to be built, first to focus energy onto a tiny spot, and then to measure the information from the particles produced, which are generally highly unstable. Paradoxically, the smaller the objects that are studied, the higher is the amount of energy needed, which is why the field is often referred to as high-energy physics (HEP). As with all fundamental research, the purpose of elementary particle physics is to seek new knowledge. This understanding can then inspire applications, for example in nanotechnology, space flight, the medical industry, or even novel types of energy production.

The largest and most versatile research center for particle physics is the European Laboratory for Particle Physics (CERN). As this thesis is being written, the commissioning of a new particle accelerator, called the Large Hadron Collider (LHC), and of its six experiments, which measure the particle collisions produced at the LHC, is being completed. The LHC has been designed to allow collision energies up to seven times higher and a collision rate of up to hundreds of times more collisions per unit of time than the currently largest operational collider, hence enabling an entirely new energy scale to be probed.

The new energy scale which is reachable with the LHC is particularly interesting, because strong theoretical arguments require it to contain information about the last missing part of the standard model of particle physics (SM), which describes the particles known thus far and the most important interactions between them. This last missing part is the breaking of the electroweak symmetry, which is assumed to account for the generation of particles' masses through the Higgs mechanism. If the Higgs mechanism is an accurate description of nature, it should manifest itself as one or more relatively heavy particles, Higgs bosons, whose properties impose considerable constraints on the possible extensions of the standard model. Hence, the existence and the properties of the Higgs boson(s) are one of the most sought after topics in particle physics.

One of the most popular candidates theories to extend the standard model is the minimal supersymmetric standard model (MSSM), which introduces a new symmetry to repair known problems in the behavior of the standard model at high energies. In this theory, five Higgs bosons are expected to exist. To discover the two heavy electrically charged and the two heavy neutral Higgs bosons of the five Higgs bosons, one of the most important, if not the most important, signature to look for is the decay of the Higgs boson to a final state involving either one or two τ leptons, respectively. The relevance of these decay channels follows from the high probability of the heavy Higgs bosons to decay to τ lepton(s) in a large portion of the permitted MSSM parameter space, and from the possibility of obtaining excellent background rejection by identifying the τ leptons based on the distinct properties of their decay. This thesis presents methods to reliably and efficiently identify τ leptons in their most favored final state, in which the τ leptons decay to hadrons. The identification of hadronically decaying τ leptons, i.e. τ jets, is applied together with other standard event selection methods to the search for the heavy neutral and charged MSSM Higgs bosons

in the $\mathrm{gg} \to \mathrm{b\bar{b}H(A)}$, $\mathrm{H, A} \to \tau\tau$ and $\mathrm{gg} \to \mathrm{tbH^{\pm}}$, $\mathrm{H^{\pm}} \to \tau^{\pm}\nu_{\tau}$ processes.

In the following, an introduction to the CERN laboratory, the LHC collider and the search for the Higgs bosons is presented. The standard model and its most favored extension candidate, the minimal supersymmetric standard model, are explained in Chapter 2. Special attention is paid to their particle contents, to the generation of the particle masses through the Higgs mechanism, and to the production and decay modes of the Higgs bosons. Chapter 2 concludes with a description of the decay modes of the τ leptons, as well as with a discussion of their helicity correlations.

The compact muon solenoid (CMS) experiment, which is one of the general-purpose detectors for measuring particle collisions at the LHC, is described in Chapter 3. The software used for simulating collision events within the CMS detector and the software used for reconstructing and analysing the physics objects from either simulated or real data are presented in Chapter 4. Furthermore, Chapter 4 is complemented with a description of the algorithms most commonly used in CMS to reconstruct basic physical objects such as jets, tracks, and vertices.

After the jets, tracks, and the primary vertex have been reconstructed in the events, it is possible to apply methods for identifying τ jets. These methods are presented in Chapter 5. Attention is also paid to the estimation of systematic uncertainties resulting from mechanical uncertainty in the position of the sensors that are used to measure the tracks of the charged particles. In addition to the τ-jet identification methods, methods commonly used in CMS to identify b jets, to measure the missing transverse energy ($\mathrm{E_T^{miss}}$), and to reconstruct top quarks are described in Chapter 5.

The methods for τ-jet identification are then applied, together with b tagging, $\mathrm{E_T^{miss}}$ measurement, and top quark reconstruction, to the full analysis of the $\mathrm{H, A} \to \tau\tau \to$ hadr. and $\mathrm{H^{\pm}} \to \tau^{\pm}\nu_{\tau} \to$ hadr. channels in Chapters 6 and 7, respectively. In Chapter 7, a further reoptimization of the τ-jet identification, including the helicity correlations, is presented for the $\mathrm{H^{\pm}} \to \tau^{\pm}\nu_{\tau} \to$ hadr. channel.

After the studies to find the heavy MSSM Higgs bosons have been presented, the thesis is concluded in Chapter 8.

1.1 CERN - the European Laboratory for Particle Physics

The keyword in the successful accomplishment of today's particle physics projects, which require an extensive set of particle accelerators and experiments to be designed, developed, operated, analyzed, and maintained in order to measure particle collisions, is internationality. By combining the efforts of several nations, a high level of research and education is feasible in particle physics without excessive economic burdens on individual nations. Additionally, no single nation has the ability to provide the technology and resources or the manpower needed for the afore-mentioned tasks.

The CERN laboratory is located in the heart of Europe, on the Franco-Swiss border near Geneva. There are about 3000 persons permanently employed by CERN, and about 6500 physicists and engineers work

there through ∼500 institutes and universities from more than 80 countries. In addition to the 20 CERN member states, which are from Europe, and which fund the operation of the laboratory from their national budgets, there are also a number of observer states, including India, Israel, Japan, Pakistan, Russia, Turkey, and the United States. Additionally, several countries are involved with CERN through industrial contracts, academic training, and participation in the experiments.

Besides the high degree of internationality, another advantage of CERN is its versatile network of particle accelerators sketched in Fig. 1.1. To provide protection from synchrotron radiation, the accelerators have been built underground. The oldest accelerator at CERN is the Proton Synchrotron (PS), which was the world's most powerful accelerator with its 28 GeV beams, when it was completed in 1959. The PS is mostly used today as a pre-accelerator and injector for the Super Proton Synchrotron (SPS). The SPS was built in 1976 and it was successfully converted into a proton-antiproton collider in 1981, which led to the discovery of the W and Z bosons in 1983 by the UA1 and UA2 experiments.

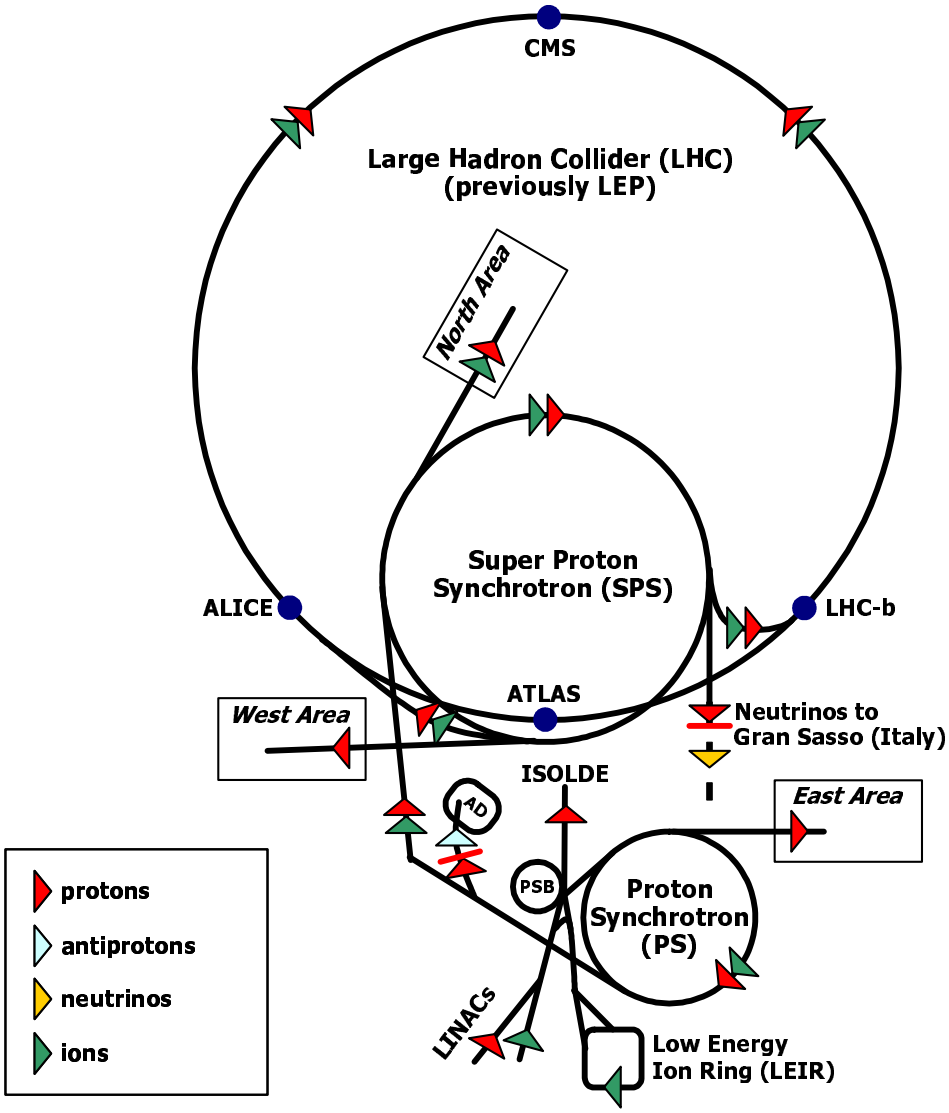

Figure 1.1: A sketch of the accelerator network at CERN and the LHC experiments (not to scale).

Today the SPS accelerates beams of protons up to energies of 450 GeV and heavy ions up to 170 GeV/nucleon for lead and indium ions. Furthermore, beams of electrons and positrons are provided. These beams are provided for various fixed target experiments and they can be used to produce secondary or tertiary beams of kaons, pions, muons, or neutrinos. The SPS also acts as a proton and heavy ion injector for the LHC. Before the LHC era, the SPS injected electrons and positrons to the Large Electron-Positron Collider (LEP), which was designed for precision measurements of the W and Z bosons. Before its shutdown in December 2000, the LEP reached an energy of 104.5 GeV per beam to search for a light Higgs boson. Nowadays it has been dismantled and the LHC occupies the tunnel constructed for the LEP. As an addition to the accelerator network, an Antimatter Decelerator (AD) has also been built to decelerate antiprotons captured from a target at PS in order to study the properties of antimatter, including the production of antiatoms such as antihydrogen.

With its unique accelerator infrastructure, CERN provides more research possibilities than any other particle physics laboratory can at the moment. Such versatility, however, also poses a challenge, because experts from all fields of physics, engineering, and information technology sciences are needed to develop and maintain all the necessary equipment to keep the system running. Furthermore, constant upgrading of the computing facilities is required in order to store and to process the vast amounts of experimental data produced by the experiments and to provide access to the data for the world-wide physics community.

1.2 LHC - the Large Hadron Collider

The Large Hadron Collider (LHC) has replaced LEP as the largest machine at CERN. Scheduled to start physics runs in 2009, the LHC will put high-energy physics well into the TeV energy scale. For proton beams, the LHC is designed to reach an energy of 7 TeV per proton and a luminosity of 10^{34} $\mathrm{cm}^{-2}\mathrm{s}^{-1}$, which represents a sevenfold increase in energy and a 30-fold increase in luminosity compared to the Tevatron accelerator at the Fermi national laboratory. The particle bunches circulating the 27 km-long underground ring built initially for the LEP are designed to be separated by only 25 ns.

The LHC also features a heavy ion program. When accelerating heavy ions, the LHC is designed to reach an energy of 2.76 TeV/nucleon and a luminosity of 10^{27} $\mathrm{cm}^{-2}\mathrm{s}^{-1}$, which will exceed the energy and the luminosity of what is currently the largest operating machine for heavy ions, the relativistic heavy ion collider (RHIC) at Brookhaven, by factors of 27 and 3, respectively. The energy density in these heavy ion collisions is expected to be large enough for studying quark-gluon plasma, which is a state of matter in which quarks are no longer bound together by the color force.

Although the LHC has been built in the same tunnel as was previously occupied by the LEP, the LHC will be able to reach higher energies than the LEP because of its choice of beam particles. The radiative losses resulting from synchrotron radiation considerably limited the maximum energy reachable with the LEP. Since the amount of synchrotron radiation is dependent on the fourth inverse power of the mass of the

particles accelerated and because the mass of protons is almost 2000 times larger than that of electrons and positrons, the radiative losses will be insignificant at the LHC compared to the LEP. However, the final states of hadron collisions are far more complex than the final states of lepton collisions, such as those of electrons and positrons, because hadrons have an inner structure, while leptons are considered point-like. Therefore, while the LEP performed highly accurate measurements on the weak force carriers and the standard model parameters, the LHC can be considered to be a machine for discovering new physics because of its ability to reach high energies.

Six experiments have been approved for the LHC to measure particle collisions. Two of these, ATLAS and CMS, are general-purpose colliding beam experiments. ALICE is a heavy ion collision experiment, whereas LHC-b is dedicated for b-quark physics with proton-proton -collisions. Furthermore, the TOTEM experiment will measure the beam luminosity and total proton-proton cross-section and study diffraction physics in the very forward directions of the CMS experiment interaction region.

1.3 The search for the Higgs boson(s)

To understand the historical background to the search for the Higgs bosons, one has to go back to the 1950s. In the 1950s, advances in accelerator technology allowed the construction of synchrotrons, which enabled scientists to study unprecedented energy scales in laboratory conditions.

Observations of K meson decays had raised suspicions that parity was not always conserved in the decays. The violation of parity in weak decays was confirmed experimentally by Wu and others in 1957 [1]. The parity non-conservation meant that the weak interactions had to be of a vector-axial vector form, which led Feynman and Gell-Mann [2] and, independently, Sudarshan and Marchak [3] to suggest that the weak interaction would be carried by charged vector bosons $\mathrm{W}^{\pm}$. Since such a formulation of the weak interaction bore similarities to electromagnetic interactions, it was tempting to try to unify these interactions with a triplet of gauge bosons ($\mathrm{W}^{+}, \gamma, \mathrm{W}^{-}$).

This attempt to formulate a unified electroweak theory was, however, found to be flawed. The first problem arose from the fact that since the weak interaction had to have a short range and since it had to be weak at low energies, its carriers, the $\mathrm{W}^{\pm}$ bosons, had to have very large masses. This feature was in contradiction to the requirement of the theory to have massless gauge bosons, such as the photon. The second problem was posed by the observation that the weak interaction did not conserve parity. This was in stark contrast with electromagnetic interaction, which was known to conserve parity.

These problems led to the suggestion of another weak interaction carrier, the Z boson. Furthermore, some mechanism had to exist to explain the breaking of the symmetry between the $\mathrm{W}^{\pm}$, Z bosons and the photon. Hence it was rather natural to try to apply spontaneous symmetry breaking to the weak interaction in a manner analogous to the spontaneous symmetry breaking in superconductivity, which had been presented by Nambu in 1960 [4]. This approach seemed, however, to fail, since the spontaneous

symmetry breaking had to produce massless Goldstone bosons according to the Goldstone theorem [5], which was in contradiction to experimental observations.

The failure only lasted for a while, since in 1964, Higgs showed that the choice of a suitable gauge in a gauge field theory would render the Goldstone theorem inapplicable [6]. With this solution, the Goldstone bosons could be "gauged away" and a new field would combine with the massless gauge bosons to form massive vector bosons. Two other groups, those of Englert and Brout [7] and Guralnik, Hagen, and Kibble [8], arrived at essentially the same conclusion independently from Higgs. This approach was later named the Englert-Brout-Higgs-Guralnik-Hagen-Kibble mechanism or, in short, the Higgs mechanism, and it is described in more detail in Sections 2.1.2 and 2.2.2.

The Higgs mechanism was able to explain the generation of masses of not only gauge bosons but also the masses of the fermions, i.e. spin-$\frac{1}{2}$ particles, of the unified electroweak theory formulated by Weinberg [9] and Salam in 1967 and 1968, respectively, as they extended the work of Glashow [10]. Furthermore, spontaneously broken gauge theories were shown to be renormalizable by 't Hooft in 1971 [11]. The electroweak theory is today known as the standard model of particle physics (SM) and it has been extended so as to also describe also the strong interactions, in addition to the electromagnetic and weak interactions.

The Higgs mechanism has an incidental side-effect, since it requires the existence of a massive scalar particle, dubbed the Higgs boson. After the $\mathrm{W}^{\pm}$ and Z bosons were discovered in 1983 [12, 13] and the top quark in 1995 [14, 15], the Higgs boson remained the only unconfirmed particle of the standard model. Hence its discovery, or experimental proof that it does not exist, has an immense impact on any theories beyond the standard model.

The Higgs mechanism is known to be flawed in the sense that it causes quadratic divergences at high energies instead of just logarithmic ones. To cancel these quadratic divergences, a minimal supersymmetry to extend the standard model was proposed by Georgi and Dimopoulos [16] in 1981. Because of this property, the minimal supersymmetric standard model (MSSM) has gained the position of the most favored theory to extend the standard model to describe the TeV energy scale, too. The price of the extension is, however, that instead of one Higgs boson, five Higgs bosons have to exist, of which one resembles the Higgs boson of the standard model. Additionally, each known particle has to have its own supersymmetric partner particle, of which none have been experimentally confirmed so far.

The mass of the Higgs boson(s) is a free parameter in the theory and hence it has been suggested that the mass of the Higgs boson(s) is too high to have been discovered by the present collider experiments. The theory does, however, supply an upper limit for the mass of the Higgs boson(s) in order for the theory to remain stable. Since this limit suggests that the mass of the Higgs boson(s) must be below the TeV energy scale, the LHC and its experiments are, after a search of four decades, finally in a position to experimentally confirm whether the Higgs boson(s) exists or not.

With the Higgs boson(s) looming within their grasp, the LHC experiments have been analysing the signatures of dozens of proposed decay channels to determine the optimal signatures to find the rare Higgs

boson signals from an almost overwhelming background of other processes. Since the Higgs mechanism generates masses to particles, the Higgs boson(s) have to have couplings with all particles. Hence, the Higgs boson(s) can decay to a staggeringly large number of different particles.

This thesis presents the simulated analyses of the $\mathrm{gg} \to \mathrm{b\bar{b}H(A)}$, $\mathrm{H, A} \to \tau\tau$ and $\mathrm{gg} \to \mathrm{tbH^{\pm}}$, $\mathrm{H^{\pm}} \to \tau^{\pm}\nu_{\tau}$ processes to find the heavy Higgs bosons in the MSSM theory with hadronic τ decays, which is the predominant τ decay mode. It turns out that these decay channels are amongst the most promising, if not the most promising decay channels, to discover the heavy MSSM Higgs bosons in a large part of the permitted MSSM parameter space as a result of distinct signature of the hadronic τ lepton decays.

The search for the Higgs boson(s) could be over within the next few years. On the other hand, the microcosm could turn out to be quite different from what anybody has managed to imagine. Either way, exciting times in the realms of high energy physics are waiting just around the corner.

2 Theory behind the Higgs bosons and τ leptons

This chapter describes the theoretical background for the work in this thesis. Section 2.1 describes the standard model of particle physics, which is a gauge field theory of the so far known fundamental particles and the electromagnetic, weak and strong interactions. Section 2.2 presents the most popular candidate theory, minimal supersymmetry between fermions and bosons, to extend the standard model beyond its current limitations. In both of these sections, special attention is paid to the Higgs mechanism, through which the masses of the different particles can be generated. The properties of the τ lepton decays, which are needed to search for the heavy MSSM Higgs bosons via the $\mathrm{gg} \to \mathrm{b\bar{b}H(A)}$, $\mathrm{H, A} \to \tau\tau$ and $\mathrm{gg} \to \mathrm{tbH^{\pm}}$, $\mathrm{H^{\pm}} \to \tau^{\pm}\nu_{\tau}$ decay modes with hadronic τ final states, are described in Section 2.3.

2.1 The standard model of particle physics

The standard model (SM) [9, 10] of particle physics was developed in the 1960s and 1970s after decades of experimental and theoretical research to describe the observations made on elementary particles and their interactions. It is a renormalizable [11] quantum field theory consistent with special relativity and it describes the interaction of spin-$\frac{1}{2}$ fermions, whose interactions are mediated by spin-1 gauge bosons.

2.1.1 The mathematical framework

The standard model, like all field theories, is formulated in terms of the Lagrangian (density) $\mathcal{L}$ from which the action S can be expressed in four-dimensional Minkowski space as

$$\mathrm{S} = \int \mathrm{d^4x}\mathcal{L}. \tag{2.1}$$

For the sake of simplicity and relevance for the topic of this thesis, the derivation of the Lagrangian of the standard model is limited in the following to the electroweak part of the standard model. The electroweak theory may then be extended to include the strong interactions. The electroweak theory is based on the the electroweak symmetry group

$$\mathrm{G_{EW}} = \mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}, \tag{2.2}$$

which unifies the electromagnetic and weak interactions [17, 18]. The fermions can be described in this theory with left-handed weak isospin doublets and right-handed weak isospin singlets with hypercharge Y defined as $\mathrm{Q} = \mathrm{I_3} + \frac{\mathrm{Y}}{2}$, where Q is the electric charge and $\mathrm{I_3}$ is the isospin. The symmetry group has 3+1 parameters and it can therefore accomodate the three vector bosons of the weak interaction and the photon of the electromagnetic interaction.

In order to construct a locally gauge invariant theory, the derivative ∂_μ of the covariant Dirac equation

$$\bar{\phi}(i\gamma^\mu\partial_\mu - \mathrm{m})\phi = 0, \tag{2.3}$$

where the Einstein summation convention $\gamma^\mu\partial_\mu = \sum_\mu \gamma_\mu\partial_\mu$ is used, and where $\phi = \phi(\mathrm{x})$ is a field, γ_μ are the Dirac matrices, and m is the mass of the field, has to be replaced with a covariant derivative. For the electroweak symmetry group, the covariant derivative can be written as

$$\mathrm{D}_\mu = \partial_\mu + \frac{\mathrm{ig}}{2}\tau_\mathrm{a}\mathrm{W}^\mathrm{a}_\mu + \frac{\mathrm{ig}'}{2}\mathrm{YB}_\mu, \tag{2.4}$$

where W_μ and B_μ are vector fields, the Pauli matrices of weak isospin τ_a and the hypercharge Y are the generators of the group, and g and g' are coupling constants.

In order to write the interaction terms of the electroweak SM Lagrangian, the currents have to be derived from equations (2.3) and (2.4). The electromagnetic current couples to the singlet fields with

$$\mathrm{J}_\mu\mathrm{B}^\mu = \bar{\psi}\frac{\mathrm{g}'}{2}\mathrm{YB}^\mu_\mathrm{a}\psi. \tag{2.5}$$

The weak currents couple only to the left-handed doublets with

$$\mathrm{J}_\mu\mathrm{W}^\mu = \bar{\psi}_\mathrm{L}\frac{\mathrm{g}}{2}\tau^\mathrm{a}\mathrm{W}^\mu_\mathrm{a}\psi_\mathrm{L}. \tag{2.6}$$

With the help of these currents and the covariant derivative, the interaction part of the electroweak SM Lagrangian can be written as

$$\begin{aligned}\mathcal{L}_\mathrm{I} \;=\; & \bar{\Psi}_\mathrm{L}\gamma^\mu(i\partial_\mu - \frac{\mathrm{g}}{2}\tau_\mathrm{a}\mathrm{W}^\mathrm{a}_\mu - \frac{\mathrm{g}'}{2}\mathrm{YB}_\mu)\Psi_\mathrm{L} + \bar{\Psi}_\mathrm{R}\gamma^\mu(i\partial_\mu - \frac{\mathrm{g}'}{2}\mathrm{YB}_\mu)\Psi_\mathrm{R} \\ & -\frac{1}{4}\mathrm{W}_{\mu\nu}\mathrm{W}^{\mu\nu} - \frac{1}{4}\mathrm{B}_{\mu\nu}\mathrm{B}^{\mu\nu},\end{aligned} \tag{2.7}$$

where Ψ_L are the left-handed lepton and quark field doublets, Ψ_R are the right-handed lepton and quark field singlets, $\mathrm{W}_{\mu\nu} = \partial_\mu\mathrm{W}_\nu - \partial_\nu\mathrm{W}_\mu$, and $\mathrm{B}_{\mu\nu} = \partial_\mu\mathrm{B}_\nu - \partial_\nu\mathrm{B}_\mu$. The first two terms describe the kinetic energy of the quarks and leptons and their couplings to the vector bosons, whereas the last two terms are the Yang-Mills terms, which describe the kinetic energies of the vector bosons and their self-interactions.

The mass terms for the particles are missing from the Lagrangian. They could be added by hand, but such maneuver would lead to an unrenormalizable theory with little predictive power. And yet, experiments have shown that most particles and the weak interaction gauge bosons have masses. To resolve this problem and to generate masses to both fermions and the gauge bosons while retaining a renormalizable gauge field theory, the Higgs mechanism can be used.

2.1.2 Generation of particle masses through the Higgs mechanism

The Higgs-Brout-Englert-Guralnik-Hagen-Kibble mechanism [6, 7, 8, 19, 20], or Higgs mechanism for short, provides perhaps the mathematically most elegant way thought of so far to generate masses to fermions and gauge bosons. A potential, called the Higgs potential, linked to a scalar Higgs field $\phi = \phi(\mathrm{x})$ can be added to the Lagrangian. Since the Lagrangian is defined as $\mathcal{L} = \mathrm{T} - \mathrm{V}$, where T are kinetic energy terms and V are potential terms, the Higgs part of the electroweak SM Lagrangian can be written as

$$\mathcal{L}_{\mathrm{Higgs}} = (\mathrm{D}_\mu\phi)^\dagger(\mathrm{D}^\mu\phi) - \mu^2\phi^\dagger\phi - \lambda(\phi^\dagger\phi)^2, \tag{2.8}$$

where D_μ is the covariant derivative of equation 2.4 and the last two terms describe the Higgs potential with μ as complex and λ as real parameter. By demanding, that $\mu^2 < 0$ and $\lambda > 0$, the potential has two minima

$$\frac{\partial \mathrm{V}}{\partial\phi^\dagger\phi} = \mu^2 + 2\lambda\phi^\dagger\phi = 0 \qquad \Rightarrow \qquad \phi^\dagger\phi = -\frac{\mu^2}{2\lambda} \equiv \frac{\mathrm{v}^2}{2}. \tag{2.9}$$

By taking a complex scalar field doublet as the field in the potential

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}, \tag{2.10}$$

where ϕ_a are real scalar fields, equation 2.9 becomes:

$$\phi_1^2 + \phi_2^2 + \phi_3^2 + \phi_4^2 = -\frac{\mu^2}{\lambda}. \tag{2.11}$$

It becomes now evident, that there exists an infinite number of non-trivial minima on a circle of radius v in the ϕ_a plane. Therefore, v is the ground state of the vacuum. It is possible to choose any set of ϕ_a satisfying equation (2.11), but once they are selected, the symmetry of equation (2.11) is spontaneously broken.

With the property of gauge invariance, the gauge can be selected in such a way, that

$$\phi_1 = \phi_2 = \phi_4 = 0, \qquad \phi_3^2 = \mathrm{v}^2. \tag{2.12}$$

The resulting Lagrangian will stay invariant, since the four fields ϕ_a are independent. Shifting the axis with v, i.e. expanding ϕ around v, results in

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \mathrm{v} + \mathrm{H}(\mathrm{x}) \end{pmatrix} \qquad \Rightarrow \qquad \phi^\dagger\phi = \frac{1}{2}(\mathrm{v} + \mathrm{H})^2 \tag{2.13}$$

for the broken symmetry, where $\mathrm{H}(\mathrm{x})$ is a scalar field called the Higgs field. By inserting the ϕ of the

broken symmetry into the Lagrangian of equation (2.8) yields

$$\begin{aligned}\mathcal{L}_{\mathrm{Higgs}} &= \left((\partial_\mu\phi^\dagger) + \frac{\mathrm{ig}}{2}\tau_\mathrm{a}\mathrm{W}^\mathrm{a}_\mu\phi^\dagger + \frac{\mathrm{ig}'}{2}\mathrm{B}_\mu\phi^\dagger\right)\cdot\left((\partial^\mu\phi) - \frac{\mathrm{ig}}{2}\tau^\mathrm{a}\mathrm{W}^\mu_\mathrm{a}\phi - \frac{\mathrm{ig}'}{2}\mathrm{B}^\mu\phi\right)\\ &\quad -\mu^2\phi^\dagger\phi - \lambda(\phi^\dagger\phi)^2\\ &= \partial_\mu\phi^\dagger\partial^\mu\phi - \frac{\mu^2}{8\lambda}\left(\mathrm{g}^2(\mathrm{W}_1)_\mu(\mathrm{W}_1)^\mu + \mathrm{g}^2(\mathrm{W}_2)_\mu(\mathrm{W}_2)^\mu\right)\\ &\quad -\frac{\mu^2}{8\lambda}(\mathrm{g}(\mathrm{W}_3)_\mu - \mathrm{g}'\mathrm{B}_\mu)(\mathrm{g}(\mathrm{W}_3)^\mu - \mathrm{g}'\mathrm{B}^\mu) + \frac{5}{2}\mu^2\mathrm{H}^2 - 2\lambda\mathrm{v}\mathrm{H}^3 - \frac{1}{2}\lambda\mathrm{H}^4\\ &\quad +\mathcal{L}_\mathrm{X},\end{aligned} \tag{2.14}$$

where $\mathcal{L}_\mathrm{X}$ are the uninteresting interaction and cross-terms. The selection of the gauge of equation (2.12) leads thus to terms also containing uneven powers of h and therefore the $\mathrm{SU}(2)$ symmetry has been spontaneously broken.

By choosing $\mathrm{W}^\pm_\mu = \frac{1}{\sqrt{2}}((\mathrm{W}_1)_\mu \mp \mathrm{i}(\mathrm{W}_2)_\mu)$ to replace the unphysical $(\mathrm{W}_1)_\mu$ and $(\mathrm{W}_2)_\mu$ fields, the second term of equation (2.14) can be written as $-\frac{\mu^2\mathrm{g}^2}{4\lambda}(\mathrm{W}^+)_\mu(\mathrm{W}^-)^\mu$. The gauge fields $(\mathrm{W}_3)_\mu$ and B_μ in equation (2.14) are not physical, but the physical neutral gauge fields A_μ and Z_μ may be constructed from them with orthogonal combinations through a rotation around a weak mixing angle θ_W:

$$\begin{pmatrix}\mathrm{A}_\mu\\ \mathrm{Z}_\mu\end{pmatrix} = \begin{pmatrix}\cos\theta_\mathrm{W} & \sin\theta_\mathrm{W}\\ -\sin\theta_\mathrm{W} & \cos\theta_\mathrm{W}\end{pmatrix}\begin{pmatrix}\mathrm{B}_\mu\\ (\mathrm{W}_3)_\mu\end{pmatrix}. \tag{2.15}$$

The weak mixing angle links also the coupling constants g and g' together with

$$\mathrm{g}\sin\theta_\mathrm{W} = \mathrm{g}'\cos\theta_\mathrm{W} \equiv e, \tag{2.16}$$

where e is the elementary electric charge. With the equations (2.15) and (2.16), the third term of the Lagrangian of the Higgs part can be written as $-\frac{\mu^2\mathrm{g}^2}{8\lambda\cos^2\theta_\mathrm{W}}\mathrm{Z}_\mu\mathrm{Z}^\mu$. With these substitutions, the Lagrangian of the Higgs part becomes

$$\mathcal{L}_\mathrm{Higgs} = \partial_\mu\phi^\dagger\partial^\mu\phi - \frac{\mu^2\mathrm{g}^2}{4\lambda}(\mathrm{W}^+)_\mu(\mathrm{W}^-)^\mu - \frac{\mu^2\mathrm{g}^2}{8\lambda\cos^2\theta_\mathrm{W}}\mathrm{Z}_\mu\mathrm{Z}^\mu + \frac{5}{2}\mu^2\mathrm{H}^2 + \mathcal{L}_\mathrm{X}. \tag{2.17}$$

With the selection $\mu^2 < 0$ and $\lambda > 0$, it is easy to see from equation (2.17), that the breaking of the symmetry has generated mass terms for the vector bosons $\mathrm{W}^\pm$ and Z (second and third terms). The mass term for the photon is missing, which is consistent with the observation that the photon is massless. The cost for the acquisition of the masses for the vector fields is, however, the introduction of a mass term for the scalar Higgs field appearing as the fourth term of equation (2.17). This mass of the Higgs boson is a free parameter to be determined experimentally.

The lepton masses are generated via the Yukawa terms of the Lagrangian

$$\mathcal{L}_\ell = \sum_{\ell \text{ generation}} \left[-\mathrm{G}_{\ell 1} \bar{\Psi}_{\ell \mathrm{L}} \phi \Psi_{\ell \mathrm{R}} - \mathrm{G}_{\ell 2} \bar{\Psi}_{\ell \mathrm{R}} \phi^\dagger \Psi_{\ell \mathrm{L}} \right], \tag{2.18}$$

where $\Psi_{\ell \mathrm{L}}$ are the left-handed lepton doublets $\begin{pmatrix} \nu_{\ell_\mathrm{L}} \\ \ell_\mathrm{L} \end{pmatrix}$, $\Psi_{\ell \mathrm{R}}$ are the right-handed lepton singlets ℓ_R, G are the Yukawa couplings, and the sum runs over the number of lepton generations. When the ϕ of the broken symmetry from equation (2.13) is inserted into equation (2.18), equation (2.18) may be rewritten as

$$\begin{aligned} \mathcal{L}_\ell &= \sum_{\ell \text{ generation}} \left[-\frac{\mathrm{G}_{\ell 1}}{\sqrt{2}} \begin{pmatrix} \bar{\nu}_{\ell_\mathrm{L}} & \bar{\ell}_\mathrm{L} \end{pmatrix} \begin{pmatrix} 0 \\ \mathrm{v}+\mathrm{H} \end{pmatrix} \ell_\mathrm{R} - \frac{\mathrm{G}_{\ell 2}}{\sqrt{2}} \bar{\ell}_\mathrm{R} \begin{pmatrix} 0 & \mathrm{v}+\mathrm{H} \end{pmatrix} \begin{pmatrix} \bar{\nu}_{\ell_\mathrm{L}} \\ \bar{\ell}_\mathrm{L} \end{pmatrix} \right] \\ &= \sum_{\ell \text{ generation}} \left[-\frac{\mathrm{G}_{\ell 1}}{\sqrt{2}} \left(\mathrm{v} \bar{\ell}_\mathrm{L} \ell_\mathrm{R} + \mathrm{H} \bar{\ell}_\mathrm{L} \ell_\mathrm{R} \right) - \frac{\mathrm{G}_{\ell 2}}{\sqrt{2}} \left(\mathrm{v} \bar{\ell}_\mathrm{R} \ell_\mathrm{L} + \mathrm{H} \bar{\ell}_\mathrm{R} \ell_\mathrm{L} \right) \right] \\ &= \sum_{\ell \text{ generation}} \left[-\frac{\mathrm{G}_\ell \mathrm{v}}{\sqrt{2}} \bar{\ell}\ell - \frac{\mathrm{G}_\ell}{\sqrt{2}} \mathrm{H} \bar{\ell}\ell \right]. \end{aligned} \tag{2.19}$$

The constant part of the first term, i.e. $\frac{\mathrm{G}_\ell \mathrm{v}}{\sqrt{2}}$, can be recognized as the mass term of each lepton and the second term couples the leptons to the Higgs boson. It should be noted, that the neutrinos are left without a mass.

To generate the quark masses, a hypercharge conjugate of ϕ has to be defined, i.e. $\phi_\mathrm{C} = -i\tau_2 \phi^*$. The ϕ_C transforms under the $\mathrm{SU}(2)$ symmetry just as ϕ, but it has the opposite hypercharge compared to ϕ. When the symmetry is broken, ϕ_C becomes $\begin{pmatrix} \mathrm{v}+\mathrm{H} \\ 0 \end{pmatrix}$. The masses of the quarks are generated via the Yukawa couplings

$$\begin{aligned} \mathcal{L}_\mathrm{q} = \sum\nolimits_{\mathrm{q} \text{ generation}} \Big(& -\mathrm{G}_{\mathrm{q}1} \bar{\Psi}_{\mathrm{qL}} \phi \Psi_{\mathrm{qR}} - \mathrm{G}_{\mathrm{q}2} \bar{\Psi}_{\mathrm{qR}} \phi^\dagger \Psi_{\mathrm{qL}} \\ & -\mathrm{G}_{\mathrm{q}1} \bar{\Psi}_{\mathrm{qL}} \phi_\mathrm{C} \Psi_{\mathrm{qR}} - \mathrm{G}_{\mathrm{q}2} \bar{\Psi}_{\mathrm{qR}} \phi_\mathrm{C}^\dagger \Psi_{\mathrm{qL}} \Big), \end{aligned} \tag{2.20}$$

where Ψ_{qL} are the left-handed quark doublets $\begin{pmatrix} \mathrm{u} \\ \mathrm{d}' \end{pmatrix}$, Ψ_{qR} are the right-handed quark singlets q_R, G_q are the Yukawa couplings, and the sum runs over the number of quark generations. When the ϕ and ϕ_C of the broken symmetry are inserted into equation (2.20) and when similar algebra is performed as for the leptons, equation (2.20) becomes

$$\mathcal{L}_\mathrm{q} = \sum_{\text{quarks}} \left(-\frac{\mathrm{G}_\mathrm{q} \mathrm{v}}{\sqrt{2}} \bar{\mathrm{q}}\mathrm{q} - \frac{\mathrm{G}_\mathrm{q}}{\sqrt{2}} \mathrm{H} \bar{\mathrm{q}}\mathrm{q} \right), \tag{2.21}$$

where the sum runs over the quark fields q. It can be seen, that the first term yields the mass of the

quarks, i.e. $\mathrm{m_q} = \frac{\mathrm{G_q v}}{\sqrt{2}}$, whereas the second term gives the coupling of each quark field to the Higgs field. It should be noted, that the strength of the coupling is directly proportional to the mass of the quark and that the Higgs couplings do not change the flavor of the quarks.

The beauty of the Higgs mechanism is that it solves the generation of the boson and fermion masses by introducing a scalar Higgs field and by spontaneously breaking the $\mathrm{SU}(2)$ symmetry. Compared to other approaches to generate the particle masses, the Higgs mechanism does not produce unphysical massless Goldstone boson fields $\theta(\mathrm{x})$ because of the choice of the gauge. The choice of a single doublet as the Higgs field has been shown above to be enough to generate the masses of the fermions and bosons of the standard model. However, the masses of the fermions and bosons are free parameters in the model and need to be measured by experiment. Consequently, also the mass of the Higgs boson is a free parameter, which needs to be determined by experiment, although rough theoretical bounds can be set to limit the mass range. The discovery of the Higgs boson is, however, not that easy, since it couples to all fermions and bosons and since the couplings are directly proportional to the masses of the fermions or bosons, as was seen above. Hence, a number of different decay channels have to be investigated in order to discover the Higgs boson, should it exist.

The full Lagrangian describing the electroweak standard model can be written as the sum of the equations (2.7), (2.8), (2.18) and (2.20) as

$$\mathcal{L}_{\mathrm{SM}} = \mathcal{L}_{\mathrm{I}} + \mathcal{L}_{\mathrm{Higgs}} + \mathcal{L}_{\ell} + \mathcal{L}_{\mathrm{q}}. \tag{2.22}$$

The strong interactions of quantum chromodynamics (QCD) can be described by the $\mathrm{SU}(3)$ color symmetry, whose eight generators are the Gell-Mann matrices $\lambda_{\mathrm{a}}^{\mathrm{GM}}$. Hence, the electroweak standard model can be extended to include the strong interactions by inserting the term $-\frac{\mathrm{g_S}}{2}\lambda_{\mathrm{a}}^{\mathrm{GM}}\mathrm{G}_{\mu}^{\mathrm{a}}$, where $\mathrm{g_S}$ is the coupling constant of the strong interactions and $\mathrm{G}_{\mu}^{\mathrm{a}}$ are the gluon fields, i.e. the carriers of the color symmetry, into the covariant derivative in the interaction part of the Lagrangian. It should be noted that this term should be allowed to appear only in conjunction with the quark fields, since the strong interaction couples only to the quarks and gluons. Since the gluon fields are massless, they should not be included into the covariant derivatives of the Higgs Lagrangian. The resulting symmetry group for the electroweak and strong theory is

$$\mathrm{G_{SM}} = \mathrm{SU}(3)_{\mathrm{C}} \otimes \mathrm{SU}(2)_{\mathrm{L}} \otimes \mathrm{U}(1)_{\mathrm{Y}}. \tag{2.23}$$

2.1.3 Particle contents of the standard model

It was established in the previous section, that the standard model of particle physics describes vector-bosons, leptons, quarks and a scalar Higgs boson which has so far not been detected. These particles are summarized in Table 2.1.

It can be seen from equations (2.7), (2.18), and (2.20) that the fermion fields come in left-handed doublets and right-handed singlets. So far, all particles belonging to three different left-handed doublets have been

discovered. These include the three generations of leptons (electron e, muon μ, tau τ and their associated neutrinos, i.e. electron neutrino ν_{e}, muon neutrino ν_{μ}, and tau neutrino ν_{τ}), and three generations of quarks (up u, down d, charm c, strange s, bottom b, and top t), all of which are spin-$\frac{1}{2}$ particles. Only left-handed neutrinos are supposed to exist, which is why they do not acquire mass as was seen in equation (2.19). The electron, muon, tau and the quarks are massive and they appear also as right-handed singlets. The electric charge of the electron, muon, and tau is -1, whereas their neutrinos are chargeless. The up-type quarks, i.e. the up, charm, and bottom quarks, have an electric charge of $+\frac{2}{3}$ and the down-type quarks, i.e. the down, strange, and top quarks, have an electric charge of $-\frac{1}{3}$. Each of the leptons and quarks have also their corresponding anti-particle, which have identical properties except for opposite electrical charge as predicted by the Dirac equation.

Type	Particles		
Left-handed leptons	$\begin{pmatrix}\nu_{\mathrm{e}}\\ \mathrm{e}\end{pmatrix}_{\mathrm{L}}$	$\begin{pmatrix}\nu_{\mu}\\ \mu\end{pmatrix}_{\mathrm{L}}$	$\begin{pmatrix}\nu_{\tau}\\ \tau\end{pmatrix}_{\mathrm{L}}$
Right-handed leptons	e_{R}	μ_{R}	τ_{R}
Left-handed quarks	$\begin{pmatrix}\mathrm{u}\\ \mathrm{d}\end{pmatrix}_{\mathrm{L}}$	$\begin{pmatrix}\mathrm{c}\\ \mathrm{s}\end{pmatrix}_{\mathrm{L}}$	$\begin{pmatrix}\mathrm{b}\\ \mathrm{t}\end{pmatrix}_{\mathrm{L}}$
Right-handed quarks	u_{R}	c_{R}	b_{R}
	d_{R}	s_{R}	t_{R}
Vector bosons and gluons	γ	$\mathrm{W}^{\pm}$,Z	G_{a}
Scalar boson	H		

Table 2.1: The particle contents of the standard model of particle physics. See text for explanation.

It can be deduced from equation (2.7), that the interactions between the fermions are carried by the gauge bosons. The photon (γ) is the massless carrier of the electromagnetic interaction and the relatively heavy $\mathrm{W}^{\pm}$ and Z^{0} are the carriers of the weak interaction. It should be noted, that the interaction terms of the standard model Lagrangian contain flavor changing terms only for the $\mathrm{W}^{\pm}$ bosons. No flavor changing neutral currents exist in the standard model. These bosons are spin-1 gauge bosons. The strong interaction is carried by eight gluons (G_{a}). The standard model is completed with the Higgs boson (H), which is a scalar particle. So far, no scalar bosons have been observed and hence, the Higgs boson remains the last missing link of the standard model.

The quarks and gluons are confined by the color gauge symmetry which allows them to exist only in colorless systems, if observed on time scales greater than those allowed by the uncertainty principle. The combinations of quarks and/or anti-quarks are known as hadrons. So far, only colorless bound states of a quark and an anti-quark (mesons) and of three quarks or three anti-quarks (baryons) are known. The ordinary matter found in nature seems to consist only of the particles of the first generation, since only the electron, proton (uud) and bound states of neutrons (udd) are stable. Muons, taus, mesons, and baryons other than the proton and neutron decay quickly, but they are routinely produced in nature or

laboratories, if enough energy is provided to produce them. The neutrinos are also stable.

Attempts to include gravitation into the standard model have so far failed, because it has not been possible to formulate a working quantum field description of gravitation. However, because gravitation is 25 orders of magnitude weaker than the weak interaction, its effect can be neglected at the currently reachable energies in high energy physics.

2.1.4 Limits on the Higgs boson mass

The Higgs field and the massive scalar Higgs boson generated by it are the last missing links in the otherwise well working standard model. To calculate the Higgs boson mass, $\mathrm{m_H}$, it is necessary to know both v and λ of the Higgs Lagrangian. Because it has not been possible to determine λ theoretically, the mass of the Higgs boson remains a free parameter and it needs thus to be determined experimentally.

High precision measurements on the standard model performed with LEP, Tevatron, and at Stanford can be used to apply constraints on the Higgs boson mass. The fit of the measured electroweak parameters yields $\mathrm{m_H} = 84^{+34}_{-26}$ GeV/c^2 [21], from which it follows that $\mathrm{m_H} < 154$ GeV/c^2 at 95% confidence level. Direct measurements at LEP have established a lower bound of $\mathrm{m_H} > 114.4$ GeV/c^2 at 95% confidence level [22].

Theoretical constraints can be used to determine further upper limits on the $\mathrm{m_H}$. Unitarity constraints of the standard model described in Ref. [23] yield an upper limit of $\mathrm{m_H} < 710$ GeV/c^2, above which the perturbative expansion ceases to be valid. Furthermore, the quartic Higgs self-coupling rises monotonically as a function of the energy scale up to which the theory is valid. If the energy scale is much larger than v, the quartic Higgs self-coupling becomes trivial, i.e. it vanishes. The condition to avoid the triviality gives an upper bound for $\mathrm{m_H}$ as a function of the cut-off energy Λ up to which the theory is valid. On the other hand, if the quartic Higgs self-coupling is non-zero but too small, the large top quark mass can cause λ to become negative. The consequence of a negative λ is that the Higgs potential would approach minus infinity and the vacuum would thus have no minima to which the symmetry could break to. The requirement, that the vacuum is stable, i.e. $\lambda > 0$, yields a lower bound for $\mathrm{m_H}$ as a function of the cut-off energy Λ. The bounds on the $\mathrm{m_H}$ and Λ resulting from the triviality and vacuum stability conditions are shown in Fig. 2.1. A more detailed description of the theoretical limits on the Higgs boson mass can be found in Ref. [18].

2.1.5 Higgs boson production and decay modes

The main Higgs boson production mechanisms at the LHC proton-proton collisions at center of mass energy of 14 TeV are the gluon-gluon fusion, vector boson fusion, $\mathrm{t\bar{t}}$ and $\mathrm{b\bar{b}}$ Higgs boson radiation, and Higgs bremsstrahlung from W or Z. The tree-level Feynman diagrams for these production processes are shown in Fig. 2.2. The cross-sections of these production modes depend on the mass of the Higgs boson

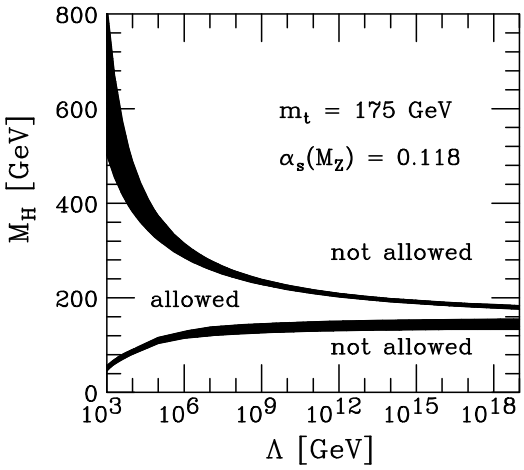

Figure 2.1: The triviality (upper) and the vacuum stability (lower) bounds on the SM Higgs boson mass $\mathrm{m_H}$ as a function of the cut-off scale Λ from Ref. [24].

as can be seen in the left plot of Fig. 2.3. The dominant production process of the Higgs boson at the LHC with $\mathrm{m_H} < 1$ TeV is the top-quark loop generated gluon-gluon fusion, which is expected to produce 10^3-10^5 Higgs bosons during the first physics run with an integrated luminosity of 10 fb^{-1} if the Higgs boson exists, as can be seen in the left plot of Fig. 2.3. A detailed discussion of the experimental and theoretical uncertainties in the Higgs boson production, such as the scale dependance and the parton distribution functions, can be found in Refs. [18, 25].

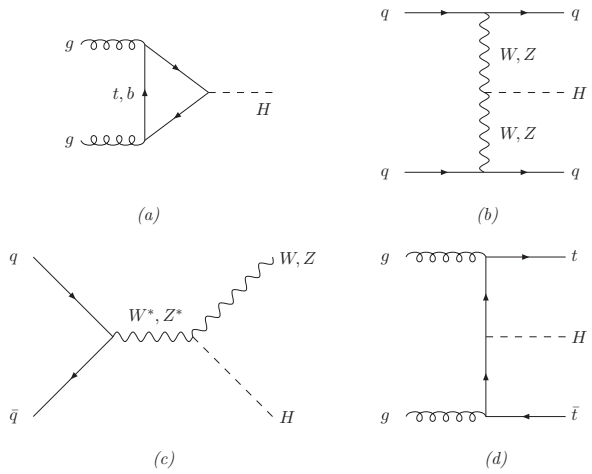

Figure 2.2: Leading order diagrams for all relevant SM Higgs boson production mechanisms: (a) gluon-gluon fusion, (b) vector boson fusion, (c) Higgs bremsstrahlung from W or Z and (d) Higgs boson radiation off top quarks.

The relevant tree-level couplings of the SM Higgs boson to the vector bosons and fermions can be identi-

fied with Feynman rules from equations (2.17), (2.19) and (2.21) as

$$\begin{cases} g_{hf\bar{f}} = i\frac{m_f}{v} \\ g_{hVV} = -2i\frac{m_V^2}{v} \end{cases}, \tag{2.24}$$

where V and f denote the vector bosons and fermions. The SM Higgs couplings are hence directly proportional to the fermion masses and quadratically proportional to the vector boson masses. Obviously, the decay to certain vector bosons or fermions has to be kinematically allowed. The branching ratios of the Higgs boson decay modes derived from the couplings are shown in the right plot of Fig. 2.3. For a Higgs boson mass of $m_H < 130$ GeV/c^2, the dominating decay channels are $H \to b\bar{b}$ and $H \to \tau\tau$ which are experimentally challenging, since they require the identification and reconstruction of multiple b jets or τ's of relatively low transverse energy from a large QCD di-jet background. Additionally, the di-τ channel requires good E_T^{miss} measurement to reconstruct the neutrinos. Hence, although the branching ratio is small, the most promising channel for the discovery of the SM Higgs boson with $m_H < 130$ GeV/c^2 is the $H \to \gamma\gamma$ because of its clean signature.

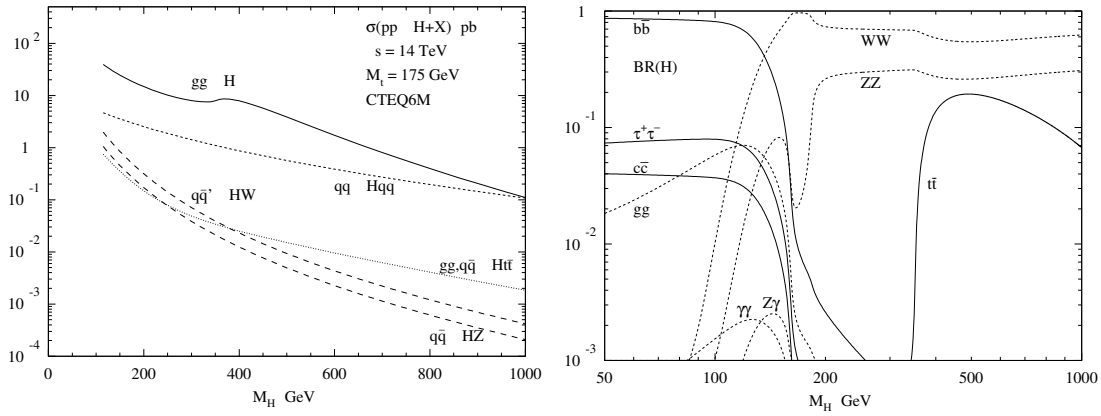

Figure 2.3: Left: The production cross-sections of the SM Higgs boson at the LHC in pb from Ref. [25]. Right: The branching ratios of the SM Higgs boson from Ref. [25]. .

For $m_H > 130$ GeV, enough energy becomes available for the Higgs boson to produce real WW and ZZ boson pairs and hence the Higgs boson decay into these channels starts to dominate. The most promising signatures of these channels for the Higgs boson discovery are given by the leptonic final states, because the amount of background events with multiple leptons and E_T^{miss} from the neutrinos is relatively small compared to the signal events. Of the leptonic final states, the $H \to ZZ \to 4\ell$ decay mode is considered to be the best channel for the Higgs boson detection and measurement over a large range of the Higgs boson mass, even though its branching ratio decreases rapidly for $m_H < 130$ GeV/c^2. For large Higgs boson masses, also the $H \to t\bar{t}$ channel opens up and may be detected via b tagging and top mass reconstruction.

The Higgs boson branching ratios rely on the running strong coupling constant α_s and the quark masses. These parameters include, however, relatively large experimental errors. The effect of varying these parameters one at a time by one standard deviation is shown in Fig. 2.4. A more detailed discussion about

the experimental and theoretical uncertainties can be found in Refs. [18, 25].

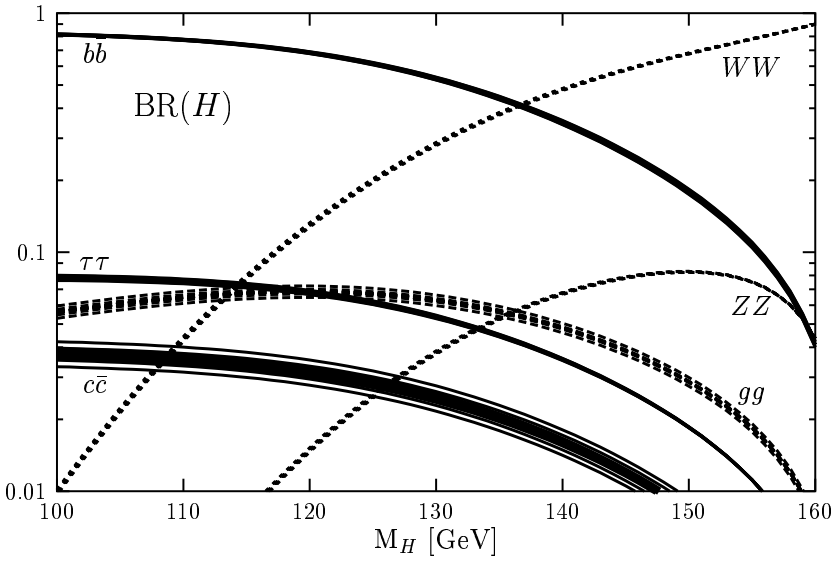

Figure 2.4: The SM Higgs boson decay branching ratios for $\mathrm{m_H} < 160$ GeV/c^2 including the uncertainties from the quark masses $\mathrm{m_t} = 178 \pm 4.3$ GeV/c^2, $\mathrm{m_b} = 4.88 \pm 0.07$ GeV/c^2 and $\mathrm{m_c} = 1.64 \pm 0.07$ GeV/c^2 as well as from $\alpha_\mathrm{s}(\mathrm{m_Z}) = 0.1172 \pm 0.002$. Taken from Ref. [18]

2.1.6 Limitations of the standard model

The extension of the standard model of electroweak and strong interactions to include gravity has not been possible so far, since no quantum theory of gravity has been formulated. Since gravity is several orders of magnitude weaker than the other known interactions, it may be neglected for practical purposes. However, since gravity is not included in the standard model, the standard model has to be considered to be fundamentally flawed. Furthermore, the gauge coupling constants of the electromagnetic, weak and strong interactions do not converge at certain energy scale in the standard model, which is not acceptable if a grand unification of the interactions is expected.

The standard model contains also other known flaws. It expects massless neutrinos, which is contradicted by observations of neutrino flavor oscillations, which suggest the neutrinos have at least a small mass. Furthermore, the standard model does not explain the difference between the fermion or vector boson masses. Additionally, although charge-parity (CP) violation in weak interactions is included naturally in the standard model, the CP-violation mechanism fails to adequately describe the matter-antimatter asymmetry.

Also the Higgs mechanism, which is the last missing link of the standard model, causes problems. The higher order terms of the perturbative expansion of the quartic Higgs field self-coupling produce quadratic mass divergences rather than the usual logarithmic divergences. Hence, a cut-off energy Λ, which defines

the energy scale up to which the standard model is valid, has to be introduced to retain the renormalizability of the standard model. However, if Λ is set to sufficiently high energies, the bare Higgs mass and its radiative corrections have to be fine-tuned to up to 16 digits to remove the quadratic divergences, which is known as the naturalness or fine-tuning problem. Since the maximum possible value of Λ depends on the mass of the Higgs boson as seen in Fig. 2.1, the observation of the Higgs boson properties would give more precise limits on the standard model and thus limit the number of candidate theories to extend the standard model. The fine-tuning problem can also be used as a practical constraint to limit the meaningful values of Λ [26]. Since the fine-tuning of two digits is required already at $\Lambda \sim 10$ TeV, it is reasonable to expect new physics to emerge at the energy scale of $\mathcal{O}(\Lambda) = 1$ TeV.

The standard model may be perceived as a behemoth of a theory, since it contains $\sim$20 parameters, which need to be determined by measurement. Although such a theory may not be the most physically elegant or meaningful description, the standard model has passed the practical tests imposed on it at LEP and at Tevatron with exceptional precision. It is because of the practical success of the standard model that most of the candidate theories to follow or to extend it use it as a basis on top of which new higher-level corrections and symmetries are built. One of the perhaps most promising of such candidate theories is supersymmetry (SUSY), which postulates an additional symmetry between fermions and bosons reducing thus the quadratical divergences in the quartic Higgs field self-couplings to logarithmic ones. It remains as the task of the new accelerators and experiments to confirm which of the new theories yields the best description of the microcosm of particles and their interactions.

2.2 MSSM - the minimal supersymmetric standard model

The perhaps most favored extension of the standard model is the minimal supersymmetric standard model (MSSM) [16, 27, 28, 29, 30]. It is based on supersymmetry [31, 32, 33], which postulates a symmetry between bosons and fermions. The price of this new symmetry is that it more than doubles the number of elementary particles that have to exist, since each particle has to have its own superpartner. Since no superpartners have been found so far, the supersymmetry must be a broken symmetry, i.e. the superpartners have to be more massive than the elementary particles of the standard model.

The benefit of these new particles is that the loops of particles and their supersymmetric partners have the ability to cancel the quadratic divergences in the Higgs field self-couplings hence solving the hierarchy problem. In order for this cure to work, it is assumed that some sign of the supersymmetric particles should manifest itself in the TeV energy scale and hence be within the reach of the LHC. One possible sign for discovering supersymmetry at the LHC could be the discovery of supersymmetric Higgs boson(s). To give masses to both normal and supersymmetric particles, the minimal Higgs sector has to consist of two Higgs doublets. Hence the MSSM boasts altogether five physical Higgs bosons instead of just one.

Supersymmetric theories are interesting if one also expects a more unified description of the interactions between the particles. They can include gravity, if the supersymmetry is made local, and they also appear

naturally in superstring theories. Furthermore, the new supersymmetric particles can allow the coupling constants of the electromagnetic, weak and strong interactions to unite [34].

In the following, the MSSM theory, its Higgs sector, and its particle contents are described in more detail.

2.2.1 Mathematical framework of the MSSM

The MSSM assumes the minimal symmetry group $\mathrm{SU(3)_C \otimes SU(2)_L \otimes U(1)_Y}$ and the minimal particle content. Since the lepton and baryon quantum numbers are not necessarily conserved in supersymmetric theories and since the they are known to be well conserved in the standard model, R-parity [35]

$$\mathrm{R} = (-1)^{2\mathrm{s}+3\mathrm{B}+\mathrm{L}}, \tag{2.25}$$

where s, B and L are the spin, baryon, and lepton quantum numbers of a particle, respectively, is assumed to be conserved in the MSSM. In the following, the MSSM Lagrangian will be derived. A more detailed derivation can be found, for example, in Ref. [30].

The MSSM assumes in addition to the four-dimensional space-time coordinates x^μ, where $\mu = 0-3$, four additional fermionic coordinates θ^1, θ^2, $\bar{\theta}^1$ and $\bar{\theta}^2$, which transform as two-component Weyl spinors. In this eight-dimensional superspace, the matter fields can be written with chiral superfields $\mathrm{S} = \mathrm{S}(\mathrm{x}, \theta, \bar{\theta})$. Hence, the left-handed doublets of leptons and their neutrinos $\mathrm{L_L}$ and quarks $\mathrm{Q_L}$ can be written as

$$\mathrm{L_L} = \mathrm{L_L}(\mathrm{x}, \theta, \bar{\theta}) = \begin{pmatrix} \nu_{\ell\mathrm{L}}(\mathrm{x}, \theta, \bar{\theta}) \\ \ell_\mathrm{L}(\mathrm{x}, \theta, \bar{\theta}) \end{pmatrix}, \quad \mathrm{Q_L} = \mathrm{Q_L}(\mathrm{x}, \theta, \bar{\theta}) = \begin{pmatrix} \mathrm{U_L}(\mathrm{x}, \theta, \bar{\theta}) \\ \mathrm{D_L}(\mathrm{x}, \theta, \bar{\theta}) \end{pmatrix}, \tag{2.26}$$

where U and D denote the up- and down-type quarks, respectively. Furthermore, the right-handed singlets of leptons $\mathrm{L_R}$, the up-type $\mathrm{U_R}$ quarks, and down-type quarks $\mathrm{D_R}$ can be written as

$$\mathrm{L_R} = \mathrm{L_R}(\mathrm{x}, \theta, \bar{\theta}) = \ell_\mathrm{R}(\mathrm{x}, \theta, \bar{\theta}), \quad \mathrm{U_R} = \mathrm{U_R}(\mathrm{x}, \theta, \bar{\theta}), \quad \mathrm{D_R} = \mathrm{D_R}(\mathrm{x}, \theta, \bar{\theta}). \tag{2.27}$$

The vector bosons of the SU(2) and U(1) symmetries can be written as $\mathrm{V^a} = \mathrm{V^a}(\mathrm{x}, \theta, \bar{\theta})$, where a = 1-3 and $\mathrm{V'} = \mathrm{V'}(\mathrm{x}, \theta, \bar{\theta})$, respectively. As in the standard model, the generators of the U(1) and SU(2) symmetries are taken to be Y and τ_a, respectively, i.e. the hypercharge and the three Pauli weak iso-spin matrices. The gluon fields, which describe the color symmetry of the strong interaction with the SU(3) symmetry group, can be written as $\mathrm{G_a} = \mathrm{G_a}(\mathrm{x}, \theta, \bar{\theta})$, where a = 1-8. The generators of the SU(3) symmetry are the eight Gell-Mann matrices $\lambda_\mathrm{a}^\mathrm{GM}$.

With the above definition of the matter fields and with the SU(3) part added to the covariant derivative of equation (2.4), the interaction part of the Lagrangian can be written as

$$\mathcal{L}_\mathrm{I}^\mathrm{MSSM} = \int \mathrm{d}^4\theta \left(\bar{\mathrm{L}}_\mathrm{L}\gamma^\mu(i\partial_\mu - \frac{\mathrm{g}}{2}\tau_\mathrm{a}\mathrm{V}_\mu^\mathrm{a} - \frac{\mathrm{g'}}{2}\mathrm{YV'}_\mu)\mathrm{L_L} + \bar{\mathrm{L}}_\mathrm{R}\gamma^\mu(i\partial_\mu - \frac{\mathrm{g'}}{2}\mathrm{YV'}_\mu)\mathrm{L_R} \right.$$

$$
\begin{aligned}
&+\bar{\mathrm{Q}}_{\mathrm{L}}\gamma^{\mu}(i\partial_{\mu}-\frac{\mathrm{g_s}}{2}\lambda_{\mathrm{a}}^{\mathrm{GM}}\mathrm{G}_{\mu}^{\mathrm{a}}-\frac{\mathrm{g}}{2}\tau_{\mathrm{a}}\mathrm{V}_{\mu}^{\mathrm{a}}-\frac{\mathrm{g}'}{2}\mathrm{YV}'_{\mu})\mathrm{Q}_{\mathrm{L}}\\
&+\bar{\mathrm{U}}_{\mathrm{R}}\gamma^{\mu}(i\partial_{\mu}-\frac{\mathrm{g_s}}{2}\lambda_{\mathrm{a}}^{\mathrm{GM}}\mathrm{G}_{\mu}^{\mathrm{a}}-\frac{\mathrm{g}'}{2}\mathrm{YV}'_{\mu})\mathrm{U}_{\mathrm{R}}\\
&+\bar{\mathrm{D}}_{\mathrm{R}}\gamma^{\mu}(i\partial_{\mu}-\frac{\mathrm{g_s}}{2}\lambda_{\mathrm{a}}^{\mathrm{GM}}\mathrm{G}_{\mu}^{\mathrm{a}}-\frac{\mathrm{g}'}{2}\mathrm{YV}'_{\mu})\mathrm{D}_{\mathrm{R}}\\
&-\frac{1}{4}\left(\mathrm{V}^{\mathrm{a}\alpha}\mathrm{V}_{\alpha}^{\mathrm{a}}\delta^2(\bar{\theta})+\mathrm{V}'_{\alpha}\mathrm{V}'^{\alpha}\delta^2(\bar{\theta})+\mathrm{G}^{\mathrm{a}\alpha}\mathrm{G}_{\alpha}^{\mathrm{a}}\delta^2(\bar{\theta})+\mathrm{h.c.}\right)\Bigg),
\end{aligned}
\tag{2.28}
$$

where g, g', and $\mathrm{g_s}$ are the coupling constants of the electromagnetic, weak, and strong interactions, respectively.

As in the standard model, the masses of the MSSM particles are expected to be generated through the spontaneous symmetry breaking of the electroweak symmetry. In the MSSM, the two Higgs doublet model is assumed. It is the most attractive model, because it is a minimal extension of the standard model and because it adds new phenomena with the fewest number of new arbitrary parameters while satisfying the theoretical constraints. The two Higgs doublets of complex scalar fields have to have an opposite hypercharge to cancel the chiral anomalies and to give masses to the up-type and down-type fermions [27, 28, 29]. Both doublets consist of complex scalar fields, which yields a total of eight degrees of freedom. Since three degrees of freedom are taken by the $\mathrm{W}^{\pm}$ and Z vector bosons as in the standard model, the remaining five degrees of freedom are taken up by five Higgs bosons.

To construct the Higgs part of the Lagrangian, the Higgs fields and the superpotential have to be defined. The Higgs doublets can be written with the superfields as

$$
\mathrm{H}_1=\mathrm{H}_1(\mathrm{x},\theta,\bar{\theta})=\begin{pmatrix}\mathrm{H}_1^1(\mathrm{x},\theta,\bar{\theta})\\ \mathrm{H}_1^2(\mathrm{x},\theta,\bar{\theta})\end{pmatrix},\qquad \mathrm{H}_2=\mathrm{H}_2(\mathrm{x},\theta,\bar{\theta})=\begin{pmatrix}\mathrm{H}_2^1(\mathrm{x},\theta,\bar{\theta})\\ \mathrm{H}_2^2(\mathrm{x},\theta,\bar{\theta})\end{pmatrix}. \tag{2.29}
$$

The most general superpotential has the form

$$
\mathrm{W}\;=\;\mu\epsilon^{ij}\mathrm{H}_1^i\mathrm{H}_1^j+\epsilon^{ij}\left(\mathrm{fH}_1^i\mathrm{L}_{\mathrm{L}}^j\mathrm{L}_{\mathrm{R}}+\mathrm{f}_1\mathrm{H}_1^i\mathrm{Q}_{\mathrm{L}}^j\mathrm{D}_{\mathrm{R}}+\mathrm{f}_2\mathrm{H}_2^i\mathrm{Q}_{\mathrm{L}}^j\mathrm{U}_{\mathrm{R}}+\mathrm{h.c.}\right), \tag{2.30}
$$

where μ is a mass parameter and ϵ^{ij} is the anti-symmetric tensor. With these identities, the Higgs part of the Lagrangian can be written as

$$
\mathcal{L}_{\mathrm{Higgs}}^{\mathrm{MSSM}}\;=\;\int\mathrm{d}^4\theta\Big((\mathrm{D}_{\mu}\mathrm{H}_1)^{\dagger}(\mathrm{D}^{\mu}\mathrm{H}_1)+(\mathrm{D}_{\mu}\mathrm{H}_2)^{\dagger}(\mathrm{D}^{\mu}\mathrm{H}_2)+\mathrm{W}\delta^2(\theta)+\bar{\mathrm{W}}\delta^2(\bar{\theta})\Big), \tag{2.31}
$$

where D_{μ} is the covariant derivative and W is the Higgs potential, which breaks the electroweak symmetry. The first two terms of equation 2.31 yield the couplings between the Higgs fields and the gauge fields.

The soft breaking of the supersymmetry, which is assumed to occur without knowledge of its mechanism,

can be introduced with the mass terms

$$\begin{aligned}\mathcal{L}_{\text{soft}}^{\text{MSSM}} \;=\; \int \mathrm{d}^4\theta\Big(&-\mathrm{m}_{\ell_\mathrm{L}}^2\,\bar{\ell}_\mathrm{L}\ell_\mathrm{L} - \mathrm{m}_{\nu_\ell}^2\,\bar{\nu}_\ell\nu_\ell - \mathrm{m}_{\ell_\mathrm{R}}^2\,\bar{\ell}_\mathrm{R}\ell_\mathrm{R}\\ &-\mathrm{m}_1^2\mathrm{H}_1^\dagger\mathrm{H}_1 - \mathrm{m}_2^2\mathrm{H}_2^\dagger\mathrm{H}_2 - \mathrm{m}_3^2\epsilon^{ij}(\mathrm{H}_i\mathrm{H}_j + \mathrm{h.c.})\\ &+\frac{1}{2}(\mathrm{M}\mathrm{V}^{\mu\mathrm{a}}\mathrm{V}_\mu^\mathrm{a} + \mathrm{M}'\mathrm{V}'^{\mu}\mathrm{V}'_{\ \mu} + \mathrm{h.c.})\Big)\delta^4(\theta,\bar{\theta})\end{aligned} \tag{2.32}$$

The Lagrangian of the MSSM can hence be written as the sum of equations (2.28), (2.31) and (2.32) as

$$\mathcal{L}^{\text{MSSM}} = \mathcal{L}_\mathrm{I}^{\text{MSSM}} + \mathcal{L}_{\text{Higgs}}^{\text{MSSM}} + \mathcal{L}_{\text{soft}}^{\text{MSSM}}. \tag{2.33}$$

To see, how the mass terms and couplings arise from the MSSM Lagrangian of equation (2.33), which was written with the superfield formalism, the Lagrangian has to be written in a component field expansion. The Higgs fields become in this expansion

$$\begin{aligned}\mathrm{H}_{1,2} = \mathrm{H}_{1,2}(\mathrm{x},\theta,\bar{\theta}) \;&=\; \begin{pmatrix}\mathrm{H}_{1,2}^1(\mathrm{x},\theta,\bar{\theta})\\ \mathrm{H}_{1,2}^2(\mathrm{x},\theta,\bar{\theta})\end{pmatrix}\\ &=\; \mathrm{H}_{1,2}(\mathrm{x}) + i\theta\sigma^\mu\bar{\theta}\partial_\mu\mathrm{H}_{1,2}(\mathrm{x}) - \frac{1}{4}\theta\theta\bar{\theta}\bar{\theta}\partial^\mu\partial_\mu\mathrm{H}_{1,2}(\mathrm{x})\\ &\quad+\sqrt{2}\theta\tilde{\mathrm{H}}_{1,2}^{(2)}(\mathrm{x}) + \frac{i}{\sqrt{2}}\theta\theta\bar{\theta}\bar{\sigma}^\mu\partial_\mu\tilde{\mathrm{H}}_{1,2}^{(2)}(\mathrm{x}) + \theta\theta\mathrm{F}_{1,2}(\mathrm{x}),\end{aligned} \tag{2.34}$$

where the $\mathrm{F}_{1,2}(\mathrm{x})$ are auxiliary fields, which vanish in the on-shell Lagrangian because of the Euler-Lagrange equation [30]. The tilde on top of the field refers to the supersymmetric partner of the Higgs fields, i.e. Higgsinos, and the superscript (2) refers to the two-component Weyl spinors. The component field expansion for the lepton and quark fields leads to a similar expression as for the Higgs fields with auxiliary fields $\mathrm{F}_\mathrm{L}(\mathrm{x})$ and $\mathrm{F}_\mathrm{R}(\mathrm{x})$ for the left-handed and right-handed contributions. For the gauge boson fields, the expansion yields

$$\mathcal{V}^\mathrm{a} = \mathcal{V}^\mathrm{a}(\mathrm{x},\theta,\bar{\theta}) \;=\; -\theta\sigma^\mu\bar{\theta}\mathcal{V}_\mu^\mathrm{a}(\mathrm{x}) + \mathrm{i}\theta\theta\bar{\theta}\bar{\lambda}^\mathrm{a}(\mathrm{x}) - \mathrm{i}\bar{\theta}\bar{\theta}\theta\lambda^\mathrm{a}(\mathrm{x}) + \frac{1}{2}\theta\theta\bar{\theta}\bar{\theta}\mathrm{D}^\mathrm{a}(\mathrm{x}), \tag{2.35}$$

where $\lambda^\mathrm{a}(\mathrm{x})$ are the two component Weyl gaugino fields, i.e. the supersymmetric partners of the gauge boson fields, and the D-fields are auxiliary fields.

The Lagrangian containing F- and D-fields is the off-shell form of the Lagrangian. The on-shell form of the Lagrangian is obtained by removing the F- and D-terms from the off-shell Lagrangian with the Euler-Lagrange equation

$$\frac{\partial\mathcal{L}}{\partial\phi} - \partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)} = 0, \tag{2.36}$$

where ϕ is any Minkowski field. The Euler-Lagrange equation is also valid for the hermitian conjugated Minkowski fields. When the Euler-Lagrange equation is applied to the F- and D-terms of the MSSM

Lagrangian, the following expressions are obtained

$$\mathrm{F}_{\mathrm{L}}^{j\dagger} = -\mathrm{f}\epsilon^{ij}\mathrm{H}_1^i\tilde{\Phi}_{\mathrm{R}} + \mathrm{h.c.} \tag{2.37}$$

$$\mathrm{F}_{\mathrm{R}}^{\dagger} = -\mathrm{f}\epsilon^{ij}\mathrm{H}_1^i\tilde{\Phi}_{\mathrm{L}}^j + \mathrm{h.c.} \tag{2.38}$$

$$\mathrm{F}_1^{i\dagger} = -\mu\epsilon^{ij}\mathrm{H}_2^j - \mathrm{f}\epsilon^{ij}\mathrm{H}_1^i\tilde{\Phi}_{\mathrm{L}}\tilde{\Phi}_{\mathrm{R}} + \mathrm{h.c.} \tag{2.39}$$

$$\mathrm{F}_2^{j\dagger} = -\mu\epsilon^{ij}\mathrm{H}_1^i + \mathrm{h.c.} \tag{2.40}$$

$$\mathrm{D}^{\mathrm{a}} = -\frac{\mathrm{g}_{\mathcal{V}}}{2}\left(\tilde{\Phi}_{\mathrm{L}}^{\dagger}\mathcal{G}^{\mathrm{a}}\Phi_{\mathrm{L}} + \tilde{\Phi}_{\mathrm{R}}^{\dagger}\mathcal{G}^{\mathrm{a}}\Phi_{\mathrm{R}} + \mathrm{H}_1^{\dagger}\mathcal{G}^{\mathrm{a}}\mathrm{H}_1 + \mathrm{H}_2^{\dagger}\mathcal{G}^{\mathrm{a}}\mathrm{H}_2 + \mathrm{h.c.}\right), \tag{2.41}$$

where $\mathrm{g}_{\mathcal{V}}$ is the coupling constant and $\mathcal{G}_{\mathrm{a}}$ are the generators of the symmetry group for the interaction in question.

By applying the component field expansion to the MSSM Lagrangian and by removing the F- and D-terms with equations (2.37)-(2.41), the interaction part of the MSSM Lagrangian can be written as

$$\begin{aligned}
\mathcal{L}_{\mathrm{I}}^{\mathrm{MSSM}} = {} & (\mathrm{D}^{\mu}\tilde{\mathrm{L}})^{\dagger}(\mathrm{D}_{\mu}\tilde{\mathrm{L}}) + (\mathrm{D}^{\mu}\tilde{\mathrm{R}})^{\dagger}(\mathrm{D}_{\mu}\tilde{\mathrm{R}}) \\
& -i\bar{\mathrm{L}}\gamma^{\mu}\mathrm{D}_{\mu}\mathrm{L} - i\bar{\mathrm{R}}\gamma^{\mu}\mathrm{D}_{\mu}\mathrm{R} \\
& -\mathrm{g}_{\mathcal{V}}^{\mathrm{a}}\left(\bar{\mathrm{L}}\mathcal{G}^{\mathrm{a}}\lambda^{\mathrm{a}}\tilde{\mathrm{L}} + \bar{\mathrm{R}}\mathcal{G}^{\mathrm{a}}\bar{\lambda}^{\mathrm{a}}\tilde{\mathrm{R}} + \mathrm{h.c.}\right) \\
& -\mathrm{g}_{\mathcal{V}}^{\mathrm{a}}\left(\bar{\lambda}^{\mathrm{a}}\gamma^{\mu}\lambda^{\mathrm{b}}\mathcal{V}_{\mu}^{\mathrm{b}} + \bar{\lambda}^{\mathrm{b}}\gamma^{\mu}\lambda^{\mathrm{a}}\mathcal{V}_{\mu}^{\mathrm{b}} - \bar{\lambda}^{\mathrm{b}}\gamma^{\mu}\lambda^{\mathrm{b}}\mathcal{V}_{\mu}^{\mathrm{a}}\right) \\
& -\frac{i}{2}\bar{\lambda}^{\mathrm{a}}\gamma^{\mu}\partial_{\mu}\lambda^{\mathrm{a}} \\
& -\frac{1}{4}\mathcal{V}^{\mathrm{a}\mu\nu}\mathcal{V}_{\mu\nu}^{\mathrm{a}},
\end{aligned} \tag{2.42}$$

where D_{μ} is the covariant derivative, $\mathrm{g}_{\mathcal{V}}$ are the coupling constants of the gauge fields, $\mathcal{G}_{\mathcal{V}}^{\mathrm{a}}$ are the generators of the symmetry group of the interaction, γ are the Dirac matrices and L and R are the left-handed doublets and right-handed singlets of leptons and quarks, respectively, as described for the standard model. The summation over the indices a and b, as well as over the particle generation, should be understood to be included in the equation. It should also be remembered, that the electromagnetic interaction couples only to fields with a hypercharge, the weak interaction couples only to the left-handed fields, and the strong interaction couples only to the (s)quarks.

By applying the component field expansion to the MSSM Lagrangian and by removing the F- and D-terms with equations (2.37)-(2.41), the Higgs part of the MSSM Lagrangian can be written as

$$\begin{aligned}
\mathcal{L}_{\mathrm{Higgs}}^{\mathrm{MSSM}} = {} & \left(\mathrm{D}^{\mu}\mathrm{H}_1\right)^{\dagger}\left(\mathrm{D}_{\mu}\mathrm{H}_1\right) + \left(\mathrm{D}^{\mu}\mathrm{H}_2\right)^{\dagger}\left(\mathrm{D}_{\mu}\mathrm{H}_2\right) \\
& -i\bar{\tilde{\mathrm{H}}}_1\gamma^{\mu}\mathrm{D}_{\mu}\tilde{\mathrm{H}}_1 - i\bar{\tilde{\mathrm{H}}}_2\gamma^{\mu}\mathrm{D}_{\mu}\tilde{\mathrm{H}}_2 \\
& +\frac{i\mathrm{g}_{\mathcal{V}}^{\mathrm{a}}}{\sqrt{2}}\left(\mathrm{H}_1^{\dagger}\mathcal{G}^{\mathrm{a}}\lambda^{\mathrm{a}}\tilde{\mathrm{H}}_1 - \bar{\tilde{\mathrm{H}}}_1\mathcal{G}^{\mathrm{a}}\bar{\lambda}^{\mathrm{a}}\mathrm{H}_1 + \mathrm{H}_2^{\dagger}\mathcal{G}^{\mathrm{a}}\lambda^{\mathrm{a}}\tilde{\mathrm{H}}_2 - \bar{\tilde{\mathrm{H}}}_2\mathcal{G}^{\mathrm{a}}\bar{\lambda}^{\mathrm{a}}\mathrm{H}_2\right) \\
& -\frac{\mu}{2}\bar{\tilde{\mathrm{H}}}_1\tilde{\mathrm{H}}_2 - \frac{\mu}{2}\bar{\tilde{\mathrm{H}}}_2\tilde{\mathrm{H}}_1 + \mu\bar{\tilde{\mathrm{H}}}_1\tilde{\mathrm{H}}_1 + \mu\bar{\tilde{\mathrm{H}}}_2\tilde{\mathrm{H}}_2
\end{aligned}$$

$$
\begin{aligned}
&-\mathrm{f}\Big(\bar{\tilde{\mathrm{H}}}_1\mathrm{L}^1\tilde{\mathrm{R}} + \bar{\tilde{\mathrm{H}}}_2\mathrm{L}^1\tilde{\mathrm{R}} - \bar{\tilde{\mathrm{H}}}_1\mathrm{L}^2\tilde{\mathrm{R}} + \bar{\mathrm{R}}\mathrm{L}^1\mathrm{H}_1^2 - \bar{\mathrm{R}}\mathrm{L}^2\mathrm{H}_1^1 \\
&\qquad +\bar{\mathrm{R}}\tilde{\mathrm{H}}_1^{\mathrm{c}}\tilde{\mathrm{L}}^1 + \bar{\mathrm{R}}\tilde{\mathrm{H}}_2^{\mathrm{c}}\tilde{\mathrm{L}}^1 - \bar{\mathrm{R}}\tilde{\mathrm{H}}_1\tilde{\mathrm{L}}^2 + \mathrm{h.c.}\Big) \\
&-\mu^2\mathrm{H}_1^\dagger\mathrm{H}_1 - \mu^2\mathrm{H}_2^\dagger\mathrm{H}_2 - \mu\mathrm{f}\left(\tilde{\mathrm{H}}_2^\dagger\tilde{\mathrm{L}}\tilde{\mathrm{R}} + \mathrm{h.c.}\right) \\
&-\mathrm{f}^2\left(\tilde{\mathrm{L}}^\dagger\tilde{\mathrm{L}}\tilde{\mathrm{R}}^\dagger\tilde{\mathrm{R}} + \mathrm{H}_1^\dagger\mathrm{H}_1\tilde{\mathrm{L}}^\dagger\tilde{\mathrm{L}} + \mathrm{H}_1^\dagger\mathrm{H}_1\tilde{\mathrm{R}}^\dagger\tilde{\mathrm{R}} + (\mathrm{H}_1\tilde{\mathrm{L}})^\dagger(\mathrm{H}_1\tilde{\mathrm{L}})\right) \\
&-\frac{\mathrm{g}^2}{2}\left(\tilde{\mathrm{L}}^\dagger\tau^{\mathrm{a}}\tilde{\mathrm{L}} + \mathrm{H}_1^\dagger\tau^{\mathrm{a}}\mathrm{H}_1 + \mathrm{H}_2^\dagger\tau^{\mathrm{a}}\mathrm{H}_2\right)^2 \\
&-\frac{\mathrm{g}^2\tan^2\theta_{\mathrm{W}}}{8}\left(\tilde{\mathrm{L}}^\dagger\tilde{\mathrm{L}} - 2\tilde{\mathrm{R}}^\dagger\tilde{\mathrm{R}} + \mathrm{H}_1^\dagger\mathrm{H}_1 - \mathrm{H}_2^\dagger\mathrm{H}_2\right)^2,
\end{aligned}
\tag{2.43}
$$

where g is the coupling of the weak interaction, θ_{W} is the weak angle, τ_{a} are the three Pauli weak isospin matrices, μ is a free parameter inherent to SUSY theories, f are constants and index c implies charge conjugation. The upper index 1 or 2 of a field should be understood as the upper (indices μ = 1,2) or lower (indices μ = 3,4) spinor of the field.

Applying the component field expansion to the last part of the MSSM Lagrangian, the soft SUSY breaking terms, yields the expression

$$
\begin{aligned}
\mathcal{L}_{\mathrm{soft}}^{\mathrm{MSSM}} \;=\; & -\mathrm{m}_{\mathrm{L}}^2\tilde{\mathrm{L}}^\dagger\tilde{\mathrm{L}} - \mathrm{m}_{\mathrm{R}}^2\tilde{\mathrm{R}}^\dagger\tilde{\mathrm{R}} - \mathrm{m}_1^2\mathrm{H}_1^\dagger\mathrm{H}_1 - \mathrm{m}_2^2\mathrm{H}_2^\dagger\mathrm{H}_2 + \mathrm{m}_3^2\epsilon^{ij}\left(\mathrm{H}_1^i\mathrm{H}_2^j + \mathrm{h.c.}\right) \\
& -\frac{1}{2}\mathrm{M}(\lambda^{\mathrm{a}}\lambda^{\mathrm{a}} + \bar{\lambda}^{\mathrm{a}}\bar{\lambda}^{\mathrm{a}}).
\end{aligned}
\tag{2.44}
$$

Combining the equations (2.42), (2.43) and (2.44) yields the total on-shell MSSM Lagrangian

$$
\mathcal{L}^{\mathrm{MSSM}} = \mathcal{L}_{\mathrm{I}}^{\mathrm{MSSM}} + \mathcal{L}_{\mathrm{Higgs}}^{\mathrm{MSSM}} + \mathcal{L}_{\mathrm{soft}}^{\mathrm{MSSM}}. \tag{2.45}
$$

It should be noted, that the MSSM Lagrangian contains the kinetic terms for all bosons and fermions. It also contains all the known standard modelcouplings, which are complemented with a plethora of new couplings amongst the supersymmetric fields and between the supersymmetric and non-supersymmetric fields.

2.2.2 Generating masses to the particles

The scalar MSSM Higgs potential can be deduced from equations (2.43) and (2.44) to be

$$
\begin{aligned}
\mathrm{V}_{\mathrm{Higgs}} \;=\; & (\mathrm{m}_1^2 + \mu^2)\mathrm{H}_1^\dagger\mathrm{H}_1 + (\mathrm{m}_2^2 + \mu^2)\mathrm{H}_2^\dagger\mathrm{H}_2 - \mathrm{m}_3^2\epsilon^{ij}\left(\mathrm{H}_1^i\mathrm{H}_2^j + \mathrm{h.c.}\right) + \\
& \frac{\mathrm{g}^2}{2}(\mathrm{H}_1^\dagger\tau^{\mathrm{a}}\mathrm{H}_1 + \mathrm{H}_2^\dagger\tau^{\mathrm{a}}\mathrm{H}_2)^2 + \frac{\mathrm{g}'^2}{8}(\mathrm{H}_1^\dagger\mathrm{H}_1 - \mathrm{H}_2^\dagger\mathrm{H}_2)^2,
\end{aligned}
\tag{2.46}
$$

where $\mathrm{g}' = \mathrm{g}\tan\theta_{\mathrm{W}}$ and where the covariant derivatives of equation (2.43) have been written explicitly. The parameters m_1^2, m_2^2, and m_3^2 can be chosen to be positive. As in the standard model, the electroweak

symmetry is broken in the MSSM via radiative corrections to generate the particle masses. The breaking of the symmetry leads to

$$\mathrm{H}_1 = \begin{pmatrix} \mathrm{v}_1 \\ 0 \end{pmatrix}, \quad \mathrm{H}_2 = \begin{pmatrix} 0 \\ \mathrm{v}_2 \end{pmatrix}, \tag{2.47}$$

where v_1 and v_2 are the two expectation values of the vacuum, which can always be chosen to be positive. With this selection, the Higgs potential can be written as

$$\mathrm{V}_{\mathrm{Higgs}} = (\mathrm{m}_1^2 + \mu^2)\mathrm{v}_1^2 + (\mathrm{m}_2^2 + \mu^2)\mathrm{v}_2^2 - 2\mathrm{m}_3^2\mathrm{v}_1\mathrm{v}_2 + \frac{1}{8}(\mathrm{g}^2 + \mathrm{g}'^2)(\mathrm{v}_1^2 - \mathrm{v}_2^2)^2. \tag{2.48}$$

Since the masses are arbitrary parameters, it is possible to choose $\mathrm{m}_{1,2}^2 + \mu^2 \rightarrow \mathrm{m}_{1,2}^2$. Hence, in order to close the potential from below, the condition $\mathrm{m}_1^2 + \mathrm{m}_2^2 - 2\mathrm{m}_3^2 \geq 0$ has to be fulfilled. The ratio of the vacuum expectation values

$$\tan\beta = \frac{\mathrm{v}_2}{\mathrm{v}_1} \tag{2.49}$$

is one of the free MSSM parameters.

The masses of the W and Z bosons can be obtained from the first two terms of equation 2.43 when equation 2.47 is applied:

$$\begin{aligned}
& (\mathrm{D}^\mu \mathrm{v}_1)^\dagger (\mathrm{D}_\mu \mathrm{v}_1) + (\mathrm{D}^\mu \mathrm{v}_2)^\dagger (\mathrm{D}_\mu \mathrm{v}_2) \\
= & \begin{pmatrix} \frac{i\mathrm{g}}{2\cos\theta_\mathrm{W}}\mathrm{v}_1\mathrm{Z}^\mu \\ \frac{i\mathrm{g}}{\sqrt{2}}\mathrm{v}_1(\mathrm{W}^-)^\mu \end{pmatrix}^\dagger \begin{pmatrix} \frac{i\mathrm{g}}{2\cos\theta_\mathrm{W}}\mathrm{v}_1\mathrm{Z}_\mu \\ \frac{i\mathrm{g}}{\sqrt{2}}\mathrm{v}_1(\mathrm{W}^-)_\mu \end{pmatrix} + \begin{pmatrix} \frac{i\mathrm{g}}{\sqrt{2}}\mathrm{v}_2(\mathrm{W}^+)^\mu \\ -\frac{i\mathrm{g}}{2\cos\theta_\mathrm{W}}\mathrm{v}_2\mathrm{Z}^\mu \end{pmatrix}^\dagger \begin{pmatrix} \frac{i\mathrm{g}}{\sqrt{2}}\mathrm{v}_2(\mathrm{W}^+)_\mu \\ -\frac{i\mathrm{g}}{2\cos\theta_\mathrm{W}}\mathrm{v}_2\mathrm{Z}_\mu \end{pmatrix} \\
= & \frac{\mathrm{g}^2}{4\cos^2\theta_\mathrm{W}}(\mathrm{v}_1^2 + \mathrm{v}_2^2)\mathrm{Z}^\mu\mathrm{Z}_\mu \\
& + \frac{\mathrm{g}^2}{4}(\mathrm{v}_1^2 + \mathrm{v}_2^2)(\mathrm{W}^+)^\mu(\mathrm{W}^-)^\mu + \frac{\mathrm{g}^2}{4}(\mathrm{v}_1^2 + \mathrm{v}_2^2)(\mathrm{W}^-)^\mu(\mathrm{W}^+)^\mu.
\end{aligned} \tag{2.50}$$

The masses of the W and Z can be hence identified as

$$\begin{cases} \mathrm{m}_\mathrm{W}^2 = \frac{\mathrm{g}^2}{2}(\mathrm{v}_1^2 + \mathrm{v}_2^2) \\ \mathrm{m}_\mathrm{Z}^2 = \frac{\mathrm{g}^2}{2\cos^2\theta_\mathrm{W}}(\mathrm{v}_1^2 + \mathrm{v}_2^2) = \frac{1}{2}(\mathrm{g}^2 + \mathrm{g}'^2)(\mathrm{v}_1^2 + \mathrm{v}_2^2) \end{cases} \tag{2.51}$$

As in the standard model, the photon remains massless in MSSM.

The lepton masses are generated from the Yukawa part of the superpotential in equation (2.43)

$$\begin{aligned}
\mathrm{f}\epsilon^{ij}\bar{\mathrm{R}}\mathrm{L}^i\mathrm{H}_1^j + \mathrm{h.c.} \rightarrow & -\mathrm{f}_\ell\mathrm{v}_1(\bar{\ell}_\mathrm{R}\ell_\mathrm{L} + \bar{\ell}_\mathrm{L}\ell_\mathrm{R}) \\
= & -\mathrm{f}_\ell\mathrm{v}_1(\bar{\ell}_\mathrm{L}\mathrm{P}_\mathrm{L}\ell_\mathrm{L} + \bar{\ell}_\mathrm{R}\mathrm{P}_\mathrm{R}\ell_\mathrm{R}) \\
= & -\mathrm{f}_\ell\mathrm{v}_1\bar{\ell}\ell,
\end{aligned} \tag{2.52}$$

where f_ℓ are constants and $\mathrm{P}_{\mathrm{L,R}}$ are the projection operators. The lepton mass terms can be identified from equation (2.52) as $\mathrm{f}_\ell\mathrm{v}_1$. The neutrinos remain massless as in the standard model. The quark masses

are generated from the same Yukawa term. The u- and d-type quarks obtain masses $\mathrm{f_q v_1}$ and $\mathrm{f_q v_2}$, respectively.

The masses of the sleptons, i.e. the supersymmetric partners of leptons, are generated from the Higgs and soft SUSY breaking parts of the MSSM Lagrangian. Since the Lagrangian contains cross-terms for the left- and right-handed sleptons, the terms have to be written in matrix form. The slepton masses are then obtained by diagonalizing their mass matrix. The mass matrix for the sleptons can be written as

$$\begin{aligned} & -\mu \mathrm{f v_2} \tilde{\ell}_\mathrm{L}^\dagger \tilde{\ell}_\mathrm{R} - \mu \mathrm{f v_2} \tilde{\ell}_\mathrm{R}^\dagger \tilde{\ell}_\mathrm{L} - \mathrm{f^2 v_1^2} \left(\tilde{\ell}_\mathrm{L}^\dagger \tilde{\ell}_\mathrm{L} + \tilde{\ell}_\mathrm{R}^\dagger \tilde{\ell}_\mathrm{R} \right) - \mathrm{m_L^2} \tilde{\ell}_\mathrm{L}^\dagger \tilde{\ell}_\mathrm{L} - \mathrm{m_R^2} \tilde{\ell}_\mathrm{R}^\dagger \tilde{\ell}_\mathrm{R} \\ = & -\begin{pmatrix} \tilde{\ell}_\mathrm{L}^\dagger & \tilde{\ell}_\mathrm{R}^\dagger \end{pmatrix} \begin{pmatrix} \mathrm{m_L^2 + f^2 v_1^2} & \mu \mathrm{f v_2} \\ \mu \mathrm{f v_2} & \mathrm{m_R^2 + f^2 v_1^2} \end{pmatrix} \begin{pmatrix} \tilde{\ell}_\mathrm{L} \\ \tilde{\ell}_\mathrm{R} \end{pmatrix}. \end{aligned} \tag{2.53}$$

The mass eigenstates, which are obtained by diagonalizing the mass matrix, are mixed with the rotating angle θ:

$$\begin{cases} \tilde{\ell}_1 = \tilde{\ell}_\mathrm{L} \cos\theta + \tilde{\ell}_\mathrm{R} \sin\theta \\ \tilde{\ell}_2 = \tilde{\ell}_\mathrm{L} \sin\theta - \tilde{\ell}_\mathrm{R} \cos\theta \end{cases}. \tag{2.54}$$

The obtained masses are

$$\begin{aligned} \mathrm{m}^2_{\tilde{\ell}_1,\tilde{\ell}_2} &= \mathrm{f^2 v_1^2} + \frac{1}{2}\left((\mathrm{m_L^2 + m_R^2}) \pm \sqrt{(\mathrm{m_L^2 + m_R^2})^2 + 4\mu^2 \mathrm{f^2 v_2^2}} \right) \\ &= \mathrm{m_\ell^2} + \frac{1}{2}\left((\mathrm{m_L^2 + m_R^2}) \pm \sqrt{(\mathrm{m_L^2 + m_R^2})^2 + 4\mu^2 \mathrm{m_\ell^2} \tan^2\beta} \right). \end{aligned} \tag{2.55}$$

Although the theory allows different masses for the left- and right-handed components, maximum mixing between the two masses, i.e. $\tilde{\mathrm{m}}_\ell^2 = \mathrm{m_L^2} = \mathrm{m_R^2}$, can be assumed. Hence,

$$\mathrm{m}^2_{\tilde{\ell}_1,\tilde{\ell}_2} = \tilde{\mathrm{m}}_\ell^2 + \mathrm{m}_\ell^2 \pm |\mu| \mathrm{m}_\ell \tan\beta. \tag{2.56}$$

The masses for the squarks are obtained similarly as the masses for the sleptons.

The superpartners of the gauge bosons γ, Z, $\mathrm{W}^\pm$, and the Higgs bosons are the bino (photino), zino, winos, and the higgsinos, respectively, which are all denoted with a tilde on top of the field. The bino, zino, and the neutral higgsinos are mixed in the Lagrangian and form neutralinos, $\tilde{\chi}^0$, whereas the Winos and the charged higgsinos form charginos, $\tilde{\chi}^\pm$, when the states are diagonalized. For the charginos, the four different fields form two independent mixings between $(\tilde{\mathrm{W}}^-, \tilde{\mathrm{H}}^-)$ and $(\tilde{\mathrm{W}}^+, \tilde{\mathrm{H}}^+)$. Hence, only two physical charginos are needed to describe the fields. By diagonalizing the mass matrix, the chargino masses can be written as

$$\begin{aligned} \mathrm{m}^2_{\tilde{\chi}_1,\tilde{\chi}_2} &= \frac{1}{2}(\mathrm{M}^2 + \mu^2) + \mathrm{m_W^2} \\ &\pm \sqrt{\frac{1}{4}(\mathrm{M}^2 - \mu^2)^2 + \mathrm{m_W^4} \cos^2(2\beta) + \mathrm{m_W^2}\left(\mathrm{M}^2 + \mu^2 + 2\mu \mathrm{M} \sin(2\beta)\right)}. \end{aligned} \tag{2.57}$$

The bino, zino, and the neutral higgsinos feature, however, more cross-terms in the MSSM Lagrangian than the Winos and the charged higgsinos. Hence, four physical states of neutralinos are needed. The mass matrix for the neutralinos can be written from the MSSM Lagrangian as

$$\mathcal{M}_{\tilde{\chi}^0} = \begin{pmatrix} \mathrm{m}_{\tilde{\mathrm{A}}} & \frac{1}{2}(\mathrm{m}_{\tilde{\mathrm{Z}}} - \mathrm{m}_{\tilde{\mathrm{A}}})\tan 2\theta_{\mathrm{W}} & 0 & 0 \\ \frac{1}{2}(\mathrm{m}_{\tilde{\mathrm{Z}}} - \mathrm{m}_{\tilde{\mathrm{A}}})\tan 2\theta_{\mathrm{W}} & \mathrm{m}_{\tilde{\mathrm{Z}}} & -\mathrm{m}_{\mathrm{Z}}\cos\beta & \mathrm{m}_{\mathrm{Z}}\sin\beta \\ 0 & -\mathrm{m}_{\mathrm{Z}}\cos\beta & 0 & -\mu \\ 0 & \mathrm{m}_{\mathrm{Z}}\sin\beta & -\mu & 0 \end{pmatrix}, \tag{2.58}$$

from which the neutralino states can be obtained through the diagonalization

$$\mathrm{Z}^{\mathrm{T}}\mathcal{M}_{\tilde{\chi}^0}\mathrm{Z}^{-1} = \mathrm{diag}(\mathrm{m}_{\chi_1^0}, \mathrm{m}_{\chi_2^0}, \mathrm{m}_{\chi_3^0}, \mathrm{m}_{\chi_4^0}). \tag{2.59}$$

The actual diagonalization of $\mathcal{M}_{\tilde{\chi}^0}$ is difficult, since the terms are interconnected. The neutralino masses are ordered by convention in ascending order with $\mathrm{m}_{\chi_1^0}$ being the lightest of the neutralino masses.

2.2.3 The Higgs boson masses

To obtain the physical Higgs boson fields and their masses, the minima of the two Higgs doublets have to be shifted in both real and imaginary parts with

$$\mathrm{H}_1 = \begin{pmatrix} \mathrm{v}_1 + \mathrm{H}_1^1 + i\mathrm{P}_1^1 \\ \mathrm{H}_1^2 \end{pmatrix}, \quad \mathrm{H}_2 = \begin{pmatrix} \mathrm{H}_2^1 \\ \mathrm{v}_2 + \mathrm{H}_2^2 + i\mathrm{P}_2^2 \end{pmatrix}. \tag{2.60}$$

The mass matrices are obtained by applying

$$\mathcal{M}_{ij}^2 = \frac{1}{2}\frac{\partial^2 \mathrm{V}_{\mathrm{Higgs}}}{\partial \mathrm{H}_i \partial \mathrm{H}_j} \tag{2.61}$$

at the limit of equation (2.60). To simplify the mass matrix expressions, the expressions obtained for m_{W} and m_{Z} in equation (2.51) and the definition of $\tan\beta$ are used. With these identities, the imaginary mass matrix can be written as

$$\mathcal{M}_{\mathrm{Im}}^2 = \begin{pmatrix} -\mathrm{m}_3^2\tan\beta & \mathrm{m}_3^2 \\ \mathrm{m}_3^2 & -\mathrm{m}_3^2\cot\beta \end{pmatrix}. \tag{2.62}$$

Since $\det(\mathcal{M}_{\mathrm{Im}}^2) = 0$, one eigenvalue of the mass matrix is zero and corresponds to the neutral Goldstone boson mass. The other eigenvalue yields the pseudoscalar Higgs boson (A) mass

$$\mathrm{m}_{\mathrm{A}}^2 = -\mathrm{m}_3^2(\tan\beta + \cot\beta) = -\frac{2\mathrm{m}_3^2}{\sin 2\beta}. \tag{2.63}$$

The neutral Goldstone boson G and the pseudoscalar Higgs boson are mixed with the angle β

$$\begin{pmatrix} \mathrm{G} \\ \mathrm{A} \end{pmatrix} = \begin{pmatrix} \cos\beta & \sin\beta \\ -\sin\beta & \cos\beta \end{pmatrix} \begin{pmatrix} \mathrm{P}_1^1 \\ \mathrm{P}_2^2 \end{pmatrix}. \tag{2.64}$$

The real part of the mass matrix can be written as

$$\mathcal{M}^2_{\mathrm{Re}} = \begin{pmatrix} -\mathrm{m}_3^2 \tan\beta + \mathrm{m}_\mathrm{Z}^2 \cos^2\beta & \mathrm{m}_3^2 - \mathrm{m}_\mathrm{Z}^2 \sin\beta\cos\beta \\ \mathrm{m}_3^2 - \mathrm{m}_\mathrm{Z}^2 \sin\beta\cos\beta & -\mathrm{m}_3^2 \cot\beta + \mathrm{m}_\mathrm{Z}^2 \sin^2\beta \end{pmatrix}. \tag{2.65}$$

With the help of the expression for m_A^2, the CP-even Higgs boson masses can be expressed from the diagonalized real mass matrix as

$$\mathrm{m}^2_{\mathrm{h,H}} = \frac{1}{2}\left(\mathrm{m}_\mathrm{A}^2 + \mathrm{m}_\mathrm{Z}^2 \mp \sqrt{(\mathrm{m}_\mathrm{A}^2 + \mathrm{m}_\mathrm{Z}^2)^2 - 4\mathrm{m}_\mathrm{A}^2\mathrm{m}_\mathrm{Z}^2\cos^2 2\beta}\right), \tag{2.66}$$

where the minus and plus signs correspond to the light (h) and heavy (H) neutral Higgs bosons, respectively. Hence, m_h is always smaller than m_Z at tree-level. The physical CP-even Higgs bosons are obtained with a rotation of angle α, which is one of the MSSM parameters

$$\begin{pmatrix} \mathrm{H} \\ \mathrm{h} \end{pmatrix} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} \mathrm{H}_1^1 \\ \mathrm{H}_2^2 \end{pmatrix}. \tag{2.67}$$

The mixing angle is given by

$$\cos 2\alpha = -\cos 2\beta \frac{\mathrm{m}_\mathrm{A}^2 - \mathrm{m}_\mathrm{Z}^2}{\mathrm{m}_\mathrm{H}^2 - \mathrm{m}_\mathrm{h}^2} \tag{2.68}$$

The mass matrix for the charged Higgs bosons can be constructed by applying equation (2.61) to the real and imaginary components of the H_1^2 and H_2^1 spinors. The mass matrix can be written as

$$\mathcal{M}^2_\pm = \frac{1}{2}\left(\mathrm{g}^2 + \frac{2\mathrm{m}_3^2}{\mathrm{v}_1\mathrm{v}_2}\right)\begin{pmatrix} \mathrm{v}_1^2 & \mathrm{v}_1\mathrm{v}_2 \\ \mathrm{v}_1\mathrm{v}_2 & \mathrm{v}_2^2 \end{pmatrix}. \tag{2.69}$$

As for the pseudo-scalar Higgs boson, the zero eigenvalue of the mass matrix is associated to a Goldstone boson. The other eigenvalue yields the charged Higgs boson masses

$$\mathrm{m}^2_{\mathrm{H}^\pm} = \frac{1}{2}\left(\mathrm{g}^2 + \frac{2\mathrm{m}_3^2}{\mathrm{v}_1\mathrm{v}_2}\right)\left(\mathrm{v}_1^2 + \mathrm{v}_2^2\right) = \mathrm{m}_\mathrm{A}^2 + \mathrm{m}_\mathrm{Z}^2. \tag{2.70}$$

Hence, $\mathrm{m}_{\mathrm{H}^\pm}$ is always greater than m_Z at tree-level. The charged Goldstone bosons $\mathrm{G}^\pm$ and the charged Higgs bosons are mixed with the angle β

$$\begin{pmatrix} \mathrm{G}^\pm \\ \mathrm{H}^\pm \end{pmatrix} = \begin{pmatrix} \cos\beta & \sin\beta \\ -\sin\beta & \cos\beta \end{pmatrix} \begin{pmatrix} \mathrm{H}_1^2 \\ \mathrm{H}_2^1 \end{pmatrix}. \tag{2.71}$$

It can be observed from the equations (2.63), (2.66), and (2.70), that the masses of the five Higgs bosons are all linked to each other. Hence, it is enough to take only one of the Higgs boson masses as a free parameter. The other free parameters are the mixing angle α, the ratio of the vacuum expectation values $\tan\beta$, and the μ-parameter and its sign.

The masses of the Higgs bosons are free parameters that need to be measured experimentally. The light Higgs boson mass $\mathrm{m_h}$ has to be less than $\sim$130 GeV/c^2, when higher order corrections are taken into account [36]. This upper bound is sensitive to the top mass value and to the mixing in the stop sector. In order to cancel the quadratic divergences in the quartic Higgs boson couplings, the masses $\mathrm{m_A}$, $\mathrm{m_H}$, and $\mathrm{m_{H^\pm}}$ need to be below $\sim$1 TeV or the Higgs self-couplings become too strong and the perturbative expansion would no longer be valid. Direct electroweak searches at LEP have excluded the MSSM Higgs bosons in the mass region below $\lesssim$ 90 GeV/c^2 at 95% confidence level [37]. Furthermore, the $\tan\beta$ parameter is constrained by $1 < \tan\beta \lesssim \frac{\mathrm{m_t}}{\mathrm{m_b}}$. The lower limit for $\tan\beta$ is based on convention and the upper limit follows from the requirement that the perturbative expansion remains valid in the theory. A more detailed discussion on the experimental and theoretical uncertainties on the MSSM Higgs boson masses can be found in Ref. [36].

2.2.4 Particle contents of the MSSM

The MSSM contains the minimal set of supersymmetric particles summarized in Table 2.2. The supersymmetric particles are denoted with a tilde on top of the symbol for the particle. The leptons and quarks have each a corresponding slepton or squark as their superpartner denoted with the tilde sign over the symbol. As in the standard model, no right-handed neutrinos exist in the MSSM.

The superpartners of the gauge bosons of the electromagnetic and weak interactions are the gauginos bino (fotino), two winos, and a zino. These form mixed states with the superpartners of the Higgs bosons, i.e. the higgsinos. The mixed states of the winos and the charged higgsinos form two physical states, which are the charginos, $\tilde{\chi}^\pm$, as described in the previous section. The bino, zino, and the neutral higgsinos form four physical states, which are the neutralinos, $\tilde{\chi}^0$. The list of supersymmetrical particles is completed with gluinos, $\tilde{\mathrm{G}}_\mathrm{a}$, which are the superpartners of gluons. Since the SU(3) color symmetry of the strong interaction has eight degrees of freedom, altogether eight gluinos are needed to describe the different color quantum number combinations.

2.2.5 The MSSM Higgs boson production and decay modes

The production mechanisms for the MSSM Higgs bosons at the 14 TeV center of mass energy proton-proton collisions at the LHC are shown in Fig. 2.5 as a function of each Higgs boson mass for $\tan\beta = 3$ and 30. The shown cross-sections include the NLO QCD corrections, but the SUSY-QCD corrections have not been applied. A detailed description of the uncertainties in the MSSM Higgs boson production cross-

Type	Particles	SUSY particles
Left-handed leptons/sleptons	$\begin{pmatrix}\nu_{\mathrm{e}}\\ \mathrm{e}\end{pmatrix}_{\mathrm{L}}, \begin{pmatrix}\nu_{\mu}\\ \mu\end{pmatrix}_{\mathrm{L}}, \begin{pmatrix}\nu_{\tau}\\ \tau\end{pmatrix}_{\mathrm{L}}$	$\begin{pmatrix}\tilde{\nu}_{\mathrm{e}}\\ \tilde{\mathrm{e}}\end{pmatrix}_{\mathrm{L}}, \begin{pmatrix}\tilde{\nu}_{\mu}\\ \tilde{\mu}\end{pmatrix}_{\mathrm{L}}, \begin{pmatrix}\tilde{\nu}_{\tau}\\ \tilde{\tau}\end{pmatrix}_{\mathrm{L}}$
Right-handed leptons/sleptons	e_{R}, μ_{R}, τ_{R}	$\tilde{\mathrm{e}}_{\mathrm{R}}$, $\tilde{\mu}_{\mathrm{R}}$, $\tilde{\tau}_{\mathrm{R}}$
Left-handed quarks/squarks	$\begin{pmatrix}\mathrm{u}\\ \mathrm{d}\end{pmatrix}_{\mathrm{L}}, \begin{pmatrix}\mathrm{c}\\ \mathrm{s}\end{pmatrix}_{\mathrm{L}}, \begin{pmatrix}\mathrm{b}\\ \mathrm{t}\end{pmatrix}_{\mathrm{L}}$	$\begin{pmatrix}\tilde{\mathrm{u}}\\ \tilde{\mathrm{d}}\end{pmatrix}_{\mathrm{L}}, \begin{pmatrix}\tilde{\mathrm{c}}\\ \tilde{\mathrm{s}}\end{pmatrix}_{\mathrm{L}}, \begin{pmatrix}\tilde{\mathrm{b}}\\ \tilde{\mathrm{t}}\end{pmatrix}_{\mathrm{L}}$
Right-handed quarks/squarks	u_{R}, c_{R}, b_{R}	$\tilde{\mathrm{u}}_{\mathrm{R}}$, $\tilde{\mathrm{c}}_{\mathrm{R}}$, $\tilde{\mathrm{b}}_{\mathrm{R}}$
	d_{R}, s_{R}, t_{R}	$\tilde{\mathrm{d}}_{\mathrm{R}}$, $\tilde{\mathrm{s}}_{\mathrm{R}}$, $\tilde{\mathrm{t}}_{\mathrm{R}}$
Vector bosons/charginos	γ, $\mathrm{W}^{\pm}$, Z	$\tilde{\chi}^{\pm}_{\mathrm{a}}$, a=1-2
Scalar bosons/neutralinos	h, H, A, $\mathrm{H}^{\pm}$	$\tilde{\chi}^{0}_{\mathrm{a}}$, a=1-4
Gluons/gluinos	G_{a}, a=1-8	$\tilde{\mathrm{G}}_{\mathrm{a}}$, a=1-8

Table 2.2: The particle contents of the minimal supersymmetric standard model of particle physics. See text for explanation.

sections can be found in Ref. [36].

It should be noted in Fig. 2.5, that the maximum allowed light Higgs boson mass is $\sim$130 GeV/c^2. The same value serves also as the lower mass limit for the heavy Higgs boson H mass. The cross-sections of the light Higgs boson production modes are found to increase as a function of $\mathrm{m_h}$ and the cross-sections of the heavy neutral Higgs boson production modes are found to decrease as a function of the their mass.

The dominating production modes of the neutral MSSM Higgs bosons are the gluon-gluon fusion ($\mathrm{gg} \rightarrow \mathrm{h/H/A}$) at low and moderate $\tan\beta$ values and Higgs radiation off b quarks ($\mathrm{pp} \rightarrow \mathrm{q\bar{q}/gg} \rightarrow \mathrm{h/H/A + b\bar{b}}$) at high $\tan\beta$ values. For low $\tan\beta$ values, also the vector boson fusion ($\mathrm{pp} \rightarrow \mathrm{qq} \rightarrow \mathrm{qq + WW/ZZ} \rightarrow \mathrm{qq + h/H}$) is relatively important in the light Higgs boson production. Other neutral Higgs boson production modes include Higgs strahlung off W and Z bosons ($\mathrm{pp} \rightarrow \mathrm{q\bar{q}} \rightarrow \mathrm{Z^*/W^*} \rightarrow \mathrm{h/HZ/W}$) and Higgs radiation of the top quarks ($\mathrm{pp} \rightarrow \mathrm{q\bar{q}/gg} \rightarrow \mathrm{h/H/A + t\bar{t}}$), but their cross-section is small compared to the dominant production mode cross-sections. The tree-level Feynman diagrams for these production processes are the same as for the SM Higgs boson in Fig. 2.2.

The production of a single charged MSSM Higgs boson is included in Fig. 2.5 as the triple dotted line. The cross-section of the single charged Higgs boson production is found to decrease rapidly as a function of $\mathrm{m_{H^{\pm}}}$ until $\sim\mathrm{m_t}$, after which the production cross-section decreases moderately as a function of $\mathrm{m_{H^{\pm}}}$. This behavior comes from the combination of two different types of production modes. For $\mathrm{m_{H^{\pm}}} \lesssim \mathrm{m_t}$, the charged Higgs boson production cross-section is dominated by the $\mathrm{gg} \rightarrow \mathrm{t\bar{t}}, \mathrm{\bar{t}} \rightarrow \mathrm{\bar{b}H^-}$ $(+\mathrm{h.c.})$ process. For $\mathrm{m_{H^{\pm}}} \gtrsim \mathrm{m_t}$, the processes $\mathrm{gb} \rightarrow \mathrm{tH^-}$ $(+\mathrm{h.c.})$ and $\mathrm{gg} \rightarrow \mathrm{t\bar{b}H^-}$ $(+\mathrm{h.c.})$ start to dominate the charged Higgs boson production cross-section. The cross-section of the latter process is by a factor $\sim$2-3 smaller than the cross-section of the $\mathrm{gb} \rightarrow \mathrm{tH^-}$ $(+\mathrm{h.c.})$ process. The tree-level Feynman

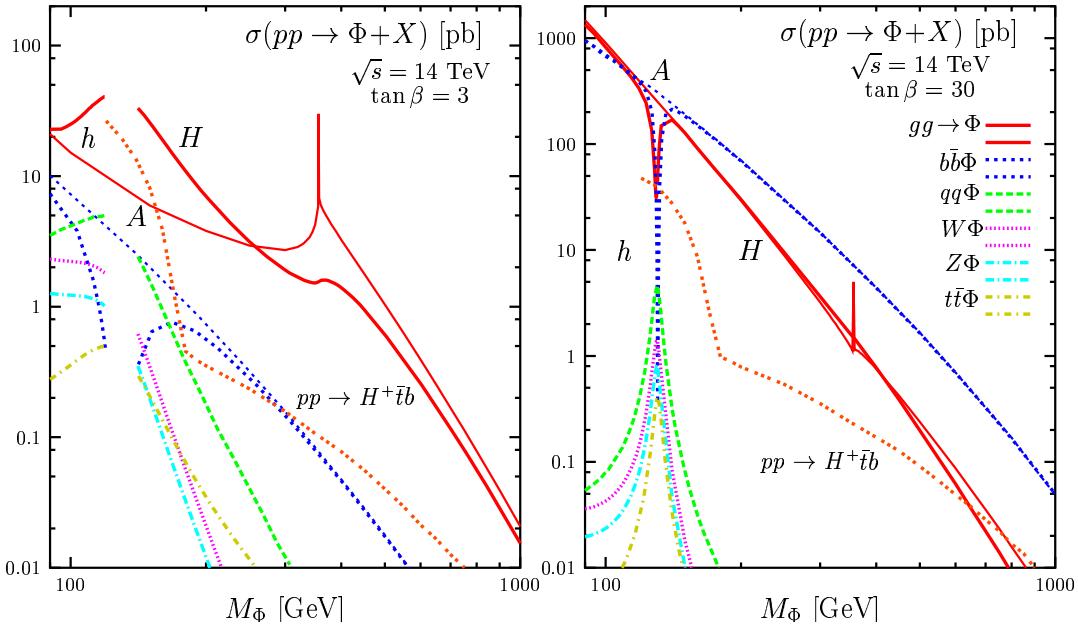

Figure 2.5: The production mode cross-sections in pb for the neutral and charged MSSM Higgs bosons at the LHC as a function of their masses for $\tan\beta = 3$ (left plot) and 30 (right plot) taken from Ref. [36]. The thick lines correspond to the h/H production, the thin lines correspond to the pseudo-scalar A production and the triple-dotted lines correspond to the charged Higgs boson production.

diagrams of the most important single charged Higgs production processes are shown in Fig. 2.6.

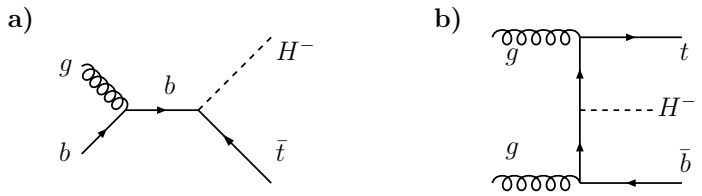

Figure 2.6: Typical leading order diagrams of charged MSSM Higgs boson production mechanisms from Ref [36]: (a) $\mathrm{gb} \to \mathrm{tH}^-$ and (b) $\mathrm{gg} \to \mathrm{t\bar{b}H}^-$.

The couplings of the MSSM Higgs bosons to (s)leptons, (s)quarks, gauge bosons, gauginos, as well as the Higgs self-interactions, can be extracted from the MSSM Lagrangian of equation (2.45) by expanding the covariant derivatives. The most important MSSM Higgs boson couplings and self-interactions are summarized in Tables 2.3-2.4 for the neutral MSSM Higgs bosons and in Table 2.5 for the charged MSSM Higgs bosons. A complete list of the MSSM Higgs boson couplings can be found in Ref. [38].

Compared to the standard model Higgs boson couplings, the MSSM Higgs boson couplings are more complex, since they are significantly affected by the $\tan\beta$ and μ parameters. To limit the MSSM parameter space, "standardized" scenarios have been proposed to be used [39]. One of these scenarios is the $\mathrm{m_h^{max}}$ scenario, which has been consistently used in this thesis. This scenario maximizes the upper bound of the $\mathrm{m_h}$ through the assumption of maximal mixing between the left- and right-handed stop

Decay type	Coupling to H	Coupling to A
$\Phi \to \mathrm{W^+W^-}$	$i\mathrm{g}\mathrm{m_W}\cos(\beta-\alpha)g^{\mu\nu}$	-
$\Phi \to \mathrm{ZZ}$	$\frac{i\mathrm{g}\mathrm{m_Z}}{\cos\theta_\mathrm{W}}\cos(\beta-\alpha)g^{\mu\nu}$	-
$\mathrm{Z} \to \mathrm{A(p')h(p)}$	-	$\frac{\mathrm{g}\cos(\beta-\alpha)}{2\mathrm{m_W}}(\mathrm{p}+\mathrm{p'})^\mu$
$\mathrm{Z} \to \mathrm{A(p')H(p)}$	-	$-\frac{\mathrm{g}\sin(\beta-\alpha)}{2\mathrm{m_W}}(\mathrm{p}+\mathrm{p'})^\mu$
$\Phi \to \mathrm{hh}$	$-\frac{i\mathrm{g}\mathrm{m_Z}}{2\cos\theta_\mathrm{W}}\Big(\ 2\sin 2\alpha\sin(\beta+\alpha)$ $-\ \cos(\beta+\alpha)\cos 2\alpha\ \Big)$	-
$\Phi \to \tilde{\chi}^-\tilde{\chi}^+$	$-\frac{ig}{\sqrt{2}}\Big(\ \mathrm{V}^*_{j1}\mathrm{U}^*_{i2}\cos\alpha\mathrm{P_L}$ $+\ \mathrm{V}^*_{j2}\mathrm{U}^*_{i1}\sin\alpha\mathrm{P_L}$ $+\ \text{complex conjugate}\ \Big)$	$-\frac{ig}{\sqrt{2}}\Big(\ \mathrm{V}^*_{j1}\mathrm{U}^*_{i2}\sin\beta\mathrm{P_L}$ $+\ \mathrm{V}^*_{j2}\mathrm{U}^*_{i1}\cos\beta\mathrm{P_L}$ $-\ \text{complex conjugate}\ \Big)$
$\Phi \to \tilde{\chi}^0\tilde{\chi}^0$	$-\frac{ig}{2}\Big(\ \mathrm{Z}^*_{j3}\mathrm{Z}^*_{i2}\cos\alpha\mathrm{P_L}$ $-\ \mathrm{Z}^*_{j3}\mathrm{Z}^*_{i1}\tan\theta_\mathrm{W}\cos\alpha\mathrm{P_L}$ $+\ \mathrm{Z}^*_{i3}\mathrm{Z}^*_{j2}\cos\alpha\mathrm{P_L}$ $-\ \mathrm{Z}^*_{i3}\mathrm{Z}^*_{j1}\tan\theta_\mathrm{W}\cos\alpha\mathrm{P_L}$ $-\ \mathrm{Z}^*_{j4}\mathrm{Z}^*_{i2}\sin\alpha\mathrm{P_L}$ $+\ \mathrm{Z}^*_{j4}\mathrm{Z}^*_{i1}\tan\theta_\mathrm{W}\sin\alpha\mathrm{P_L}$ $-\ \mathrm{Z}^*_{i4}\mathrm{Z}^*_{j2}\sin\alpha\mathrm{P_L}$ $+\ \mathrm{Z}^*_{i4}\mathrm{Z}^*_{j1}\tan\theta_\mathrm{W}\sin\alpha\mathrm{P_L}$ $+\ \text{complex conjugate}\ \Big)$	$-\frac{g}{2}\Big(\ \mathrm{Z}^*_{j3}\mathrm{Z}^*_{i2}\sin\beta\mathrm{P_L}$ $-\ \mathrm{Z}^*_{j3}\mathrm{Z}^*_{i1}\tan\theta_\mathrm{W}\sin\beta\mathrm{P_L}$ $+\ \mathrm{Z}^*_{i3}\mathrm{Z}^*_{j2}\sin\beta\mathrm{P_L}$ $-\ \mathrm{Z}^*_{i3}\mathrm{Z}^*_{j1}\tan\theta_\mathrm{W}\sin\beta\mathrm{P_L}$ $-\ \mathrm{Z}^*_{j4}\mathrm{Z}^*_{i2}\cos\beta\mathrm{P_L}$ $+\ \mathrm{Z}^*_{j4}\mathrm{Z}^*_{i1}\tan\theta_\mathrm{W}\cos\beta\mathrm{P_L}$ $-\ \mathrm{Z}^*_{i4}\mathrm{Z}^*_{j2}\cos\beta\mathrm{P_L}$ $+\ \mathrm{Z}^*_{i4}\mathrm{Z}^*_{j1}\tan\theta_\mathrm{W}\cos\beta\mathrm{P_L}$ $+\ \text{complex conjugate}\ \Big)$

Table 2.3: The most important couplings of the heavy neutral MSSM Higgs bosons H and A to gauge bosons, gauginos and other MSSM Higgs bosons from Ref. [38]. The couplings of the light MSSM Higgs boson h are obtained from those of the H by swapping $\sin\alpha \to \cos\alpha$ and $\cos\alpha \to -\sin\alpha$. The symbols $\mathrm{P_{L,R}} = \frac{1\mp\gamma_5}{2}$ are the projection operators, V_{ij} and U_{ij} are the 2×2 unitary matrices describing the charginos and Z_{ij} is the 4×4 matrix corresponding to the neutralinos.

eigenstates. The scenario applies the constraints $\mathrm{m_{SUSY}}$ = 1 TeV, gaugino mass $\mathrm{m_2}$ = 200 GeV, gluino mass $\mathrm{m_{\tilde{G}}}$ = 800 GeV, and $\mathrm{X_t} = \mathrm{A_t} - \mu\cot\beta$ = 2 TeV, where $\mathrm{A_t}$ is the trilinear Higgs-stop coupling, to the parameter space to maximize the $\mathrm{m_h}$. Maximum mixing of the sleptons and squarks is also assumed similarly as described in equation (2.56). The μ parameter and the top quark mass have been chosen to be μ = 200 GeV/c^2 and $\mathrm{m_t}$ = 175 GeV/c^2 in this thesis, respectively, unless otherwise stated. The $\mathrm{m_h^{max}}$ scenario fixes the MSSM theory up to the $\tan\beta$ and the mass of one of the Higgs bosons. These remaining free parameters are studied by varying them.

With the constraints of the $\mathrm{m_h^{max}}$ scenario, the MSSM Higgs boson couplings can be used to calculate the branching fractions of the MSSM Higgs bosons. The obtained branching ratios are shown in Figs. 2.7-2.10 for $\tan\beta$ = 3 and 30. The calculation was done with the FeynHiggs program version 2.6.5 [40, 41, 42], which takes into account NNLO corrections to the couplings. A detailed description of the uncertainties in the MSSM Higgs boson branching ratios can be found in Ref. [36].

The branching ratios of the light MSSM Higgs boson h are shown in Fig. 2.7 for $\tan\beta$ = 3 and 30 and

Decay type	Coupling to H	Coupling to A
$\Phi \to \mathrm{u\bar{u}}$	$\frac{-i\mathrm{g}\mathrm{m_u}\sin\alpha}{2\mathrm{m_W}\sin\beta}$	$\frac{-\mathrm{g}\mathrm{m_u}\cot\beta}{2\mathrm{m_W}}\gamma_5$
$\Phi \to \mathrm{d\bar{d}}$	$\frac{-i\mathrm{g}\mathrm{m_d}\cos\alpha}{2\mathrm{m_W}\cos\beta}$	$\frac{-\mathrm{g}\mathrm{m_d}\tan\beta}{2\mathrm{m_W}}\gamma_5$
$\Phi \to \ell^-\ell^+$	$\frac{-i\mathrm{g}\mathrm{m}_\ell\cos\alpha}{2\mathrm{m_W}\cos\beta}$	$\frac{-\mathrm{g}\mathrm{m}_\ell\tan\beta}{2\mathrm{m_W}}\gamma_5$
$\Phi \to \tilde{\mathrm{u}}_\mathrm{L}\bar{\tilde{\mathrm{u}}}_\mathrm{L}$	$-\frac{i\mathrm{g}\mathrm{m_Z}}{2\cos\theta_\mathrm{W}}\cos(\alpha+\beta)$ $+\frac{i\mathrm{g}\mathrm{m_Z}}{\cos\theta_\mathrm{W}}\mathrm{q_u}\sin^2\theta_\mathrm{W}\cos(\alpha+\beta)$ $-\frac{i\mathrm{g}\mathrm{m_u^2}}{\mathrm{m_W}}\frac{\sin\alpha}{\sin\beta}$	-
$\Phi \to \tilde{\mathrm{u}}_\mathrm{R}\bar{\tilde{\mathrm{u}}}_\mathrm{R}$	$-\frac{i\mathrm{g}\mathrm{m_Z}}{\cos\theta_\mathrm{W}}\mathrm{q_u}\sin^2\theta_\mathrm{W}\cos(\alpha+\beta)$ $-\frac{i\mathrm{g}\mathrm{m_u^2}}{\mathrm{m_W}}\frac{\sin\alpha}{\sin\beta}$	-
$\Phi \to \tilde{\mathrm{u}}_\mathrm{L}\bar{\tilde{\mathrm{u}}}_\mathrm{R}$	$-\frac{i\mathrm{g}\mathrm{m_u}}{2\mathrm{m_W}\sin\beta}(\mu\cos\alpha+\mathrm{A_u}\sin\alpha)$	$-\frac{\mathrm{g}\mathrm{m_u}}{2\mathrm{m_W}}(\mu-\mathrm{A_u}\cot\beta)$
$\Phi \to \tilde{\mathrm{d}}_\mathrm{L}\bar{\tilde{\mathrm{d}}}_\mathrm{L}$	$\frac{i\mathrm{g}\mathrm{m_Z}}{2\cos\theta_\mathrm{W}}\cos(\alpha+\beta)$ $+\frac{i\mathrm{g}\mathrm{m_Z}}{\cos\theta_\mathrm{W}}\mathrm{q_d}\sin^2\theta_\mathrm{W}\cos(\alpha+\beta)$ $-\frac{i\mathrm{g}\mathrm{m_d^2}}{\mathrm{m_W}}\frac{\cos\alpha}{\cos\beta}$	-
$\Phi \to \tilde{\mathrm{d}}_\mathrm{R}\bar{\tilde{\mathrm{d}}}_\mathrm{R}$	$-\frac{i\mathrm{g}\mathrm{m_Z}}{\cos\theta_\mathrm{W}}\mathrm{q_d}\sin^2\theta_\mathrm{W}\cos(\alpha+\beta)$ $-\frac{i\mathrm{g}\mathrm{m_d^2}}{\mathrm{m_W}}\frac{\cos\alpha}{\cos\beta}$	-
$\Phi \to \tilde{\mathrm{d}}_\mathrm{L}\bar{\tilde{\mathrm{d}}}_\mathrm{R}$	$-\frac{i\mathrm{g}\mathrm{m_d}}{2\mathrm{m_W}\cos\beta}(\mu\sin\alpha+\mathrm{A_d}\cos\alpha)$	$-\frac{\mathrm{g}\mathrm{m_d}}{2\mathrm{m_W}}(\mu-\mathrm{A_d}\tan\beta)$
$\Phi \to \tilde{\ell}_\mathrm{L,R}\bar{\tilde{\ell}}_\mathrm{L,R}$	$\Phi \to \tilde{\mathrm{d}}_\mathrm{L,R}\bar{\tilde{\mathrm{d}}}_\mathrm{L,R}\mid_{\mathrm{d}\to\ell}$	$\Phi \to \tilde{\mathrm{d}}_\mathrm{L,R}\bar{\tilde{\mathrm{d}}}_\mathrm{L,R}\mid_{\mathrm{d}\to\ell}$

Table 2.4: The most important couplings of the heavy neutral MSSM Higgs bosons H and A to (s)quarks and (s)leptons from Ref. [38]. The couplings of the light MSSM Higgs boson h are obtained from those of the H by swapping $\sin\alpha \to \cos\alpha$ and $\cos\alpha \to -\sin\alpha$. The symbols $\mathrm{A_x}$ denote trilinear couplings and the $\mathrm{q_x}$ symbols denote the electric charge of the particle in question.

μ = 200 GeV/c^2. The branching ratios are found to be similar to those of a light SM Higgs boson. The $\mathrm{b\bar{b}}$ decay mode is found to dominate with $\tau\tau$ being the subleading decay mode for both low and high $\tan\beta$ values. The allowed $\mathrm{m_h}$ mass range limits the h decay modes to the heavier particles such as to $\mathrm{t\bar{t}}$, to gauge boson pairs, and to supersymmetric particles. The gauge boson pair decay modes become significant only close to the maximum allowed $\mathrm{m_h}$ value. The discovery channels for h are largely the same as for the SM Higgs boson in the parameter space which has not been already excluded by LEP.

The branching ratios of H are shown in Fig. 2.8 for $\tan\beta$ = 3 and 30 and μ = 200 GeV/c^2. For low $\tan\beta$ values and for $\mathrm{m_H} < \mathrm{m_t}$, where H tends to be SM-like, the decay modes to the weak gauge boson pairs WW and ZZ and to $\mathrm{b\bar{b}}$ dominate. The $\mathrm{H} \to \mathrm{hh}$ decay mode is also sizeable. For $\mathrm{m_H} > 300$ GeV/c^2, the neutralino and chargino channels are the dominating H decay modes. Once the $\mathrm{t\bar{t}}$ decay mode opens up, it is found to have a branching ratio comparable to the neutralino and chargino channels. The $\tau\tau$ decay mode has a small branching ratio of $\sim 10^{-3}$-5×10^{-2} for $\tan\beta$ = 3. The decay modes involving neutralinos and charginos are experimentally challenging as a result of multiple sources of missing energy. Furthermore, the $\mathrm{t\bar{t}}$ and $\mathrm{b\bar{b}}$ channels are experimentally difficult because of the large QCD multi-jet background.

For high $\tan\beta$ values, the H couplings involving d-type quarks as well as electrons, muons and τ leptons

Decay type	Coupling to H^-
$\mathrm{W}^- \to \mathrm{H}^-(\mathrm{p}')\mathrm{h}(\mathrm{p})$	$\frac{ig}{2}\cos(\beta-\alpha)(\mathrm{p}+\mathrm{p}')^\mu$
$\mathrm{W}^- \to \mathrm{H}^-(\mathrm{p}')\mathrm{H}(\mathrm{p})$	$-\frac{ig}{2}\sin(\beta-\alpha)(\mathrm{p}+\mathrm{p}')^\mu$
$\mathrm{W}^- \to \mathrm{H}^-(\mathrm{p}')\mathrm{A}(\mathrm{p})$	$\frac{g}{2}(\mathrm{p}+\mathrm{p}')^\mu$
$\mathrm{H}^- \to \mathrm{d}\bar{\mathrm{u}}$	$\frac{-ig}{\sqrt{2}\mathrm{m_W}}\Big(\mathrm{m_d}\tan\beta\mathrm{P_R}+\mathrm{m_u}\cot\beta\mathrm{P_L}\Big)$
$\mathrm{H}^- \to \ell\bar{\nu}_\ell$	$\frac{-ig}{\sqrt{2}\mathrm{m_W}}\Big(\mathrm{m}_\ell\tan\beta\mathrm{P_R}\Big)$
$\mathrm{H}^- \to \tilde{\chi}^-\tilde{\chi}^0$	$\left\{\begin{array}{l} -ig\Big(\ (\mathrm{Z}_{i4}\mathrm{V}_{j1}+\frac{1}{\sqrt{2}}(\mathrm{Z}_{i2}+\mathrm{Z}_{i1}\tan\theta_\mathrm{W})\mathrm{V}_{j2})\cos\beta\mathrm{P_L} \\ \quad +(\mathrm{Z}_{i3}\mathrm{U}_{j1}+\frac{1}{\sqrt{2}}(\mathrm{Z}_{i2}+\mathrm{Z}_{i1}\tan\theta_\mathrm{W})\mathrm{U}_{j2})\cos\beta\mathrm{P_R}\ \Big) \end{array}\right.$
$\mathrm{H}^- \to \tilde{\mathrm{d}}_\mathrm{L}\bar{\tilde{\mathrm{u}}}_\mathrm{L}$	$-\frac{ig\mathrm{m_W}}{\sqrt{2}}\Big(\sin 2\beta-\frac{\mathrm{m_d^2}\tan\beta+\mathrm{m_u^2}\cot\beta}{\mathrm{m_W^2}}\ \Big)$
$\mathrm{H}^- \to \tilde{\mathrm{d}}_\mathrm{R}\bar{\tilde{\mathrm{u}}}_\mathrm{R}$	$\frac{ig\mathrm{m_u}\mathrm{m_d}}{\sqrt{2}\mathrm{m_W}}\Big(\tan\beta+\cot\beta\Big)$
$\mathrm{H}^- \to \tilde{\mathrm{d}}_\mathrm{R}\bar{\tilde{\mathrm{u}}}_\mathrm{L}$	$\frac{-ig\mathrm{m_u}}{\sqrt{2}\mathrm{m_W}}\Big(\mu-\mathrm{A_d}\tan\beta\Big)$
$\mathrm{H}^- \to \tilde{\mathrm{d}}_\mathrm{L}\bar{\tilde{\mathrm{u}}}_\mathrm{R}$	$\frac{-ig\mathrm{m_u}}{\sqrt{2}\mathrm{m_W}}\Big(\mu-\mathrm{A_u}\cot\beta\Big)$
$\mathrm{H}^- \to \tilde{\mathrm{d}}_\mathrm{L}\bar{\tilde{\mathrm{u}}}_\mathrm{L}$	$\mathrm{H}^- \to \tilde{\mathrm{d}}_{\mathrm{L,R}}\bar{\tilde{\mathrm{u}}}_\mathrm{L}\mid_{\mathrm{u}\to\nu_\ell,\mathrm{d}\to\ell}$

Table 2.5: The most important couplings of the charged MSSM Higgs bosons $\mathrm{H}^\pm$ from Ref. [38]. The symbols $\mathrm{P_{L,R}} = \frac{1\mp\gamma_5}{2}$ are the projection operators, V_{ij} and U_{ij} are the 2×2 unitary matrices describing the charginos and Z_{ij} is the 4×4 matrix corresponding to the neutralinos. The symbols $\mathrm{A_x}$ denote trilinear couplings and the $\mathrm{q_x}$ symbols denote the electric charge of the particle in question.

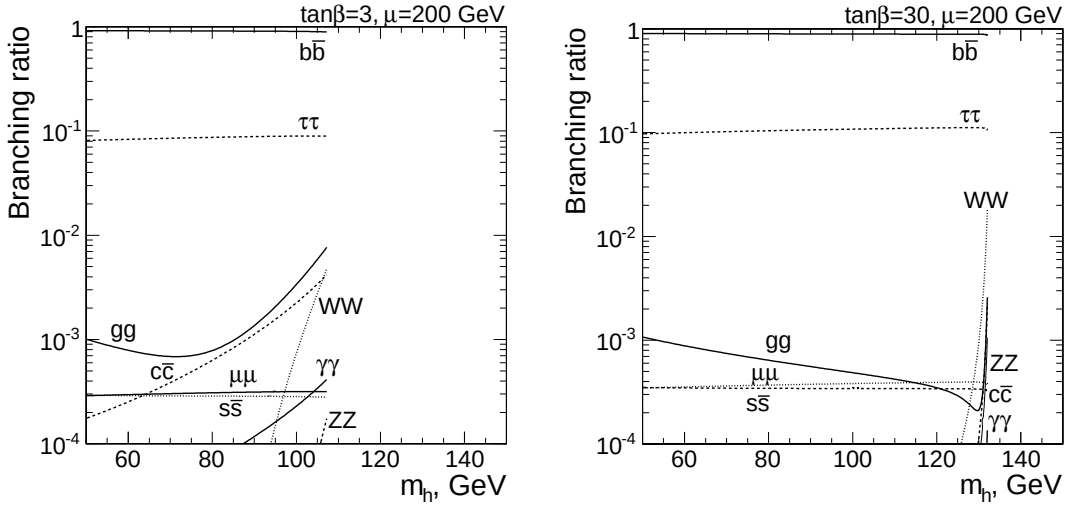

Figure 2.7: Branching ratios of the MSSM light Higgs boson h as a function of $\mathrm{m_h}$ for $\tan\beta$ = 3 (left plot) and $\tan\beta$ = 30 (right plot).

receive a boost from the $\frac{1}{\cos\beta}$ factor in the couplings. Therefore, $\mathrm{b\bar{b}}$ becomes the dominating decay mode of H over the whole $\mathrm{m_H}$ range for high $\tan\beta$ values. The H decays to neutralinos and charginos are the subleading decay modes. The experimentally most promising decay mode for the H discovery at high $\tan\beta$ values is the $\tau\tau$ decay mode, which has a significant branching ratio over the whole $\mathrm{m_H}$ mass range as a result of the $\frac{1}{\cos\beta}$ factor. The $\mathrm{H}\to\mu\mu$ decay mode is also promising because of the clean

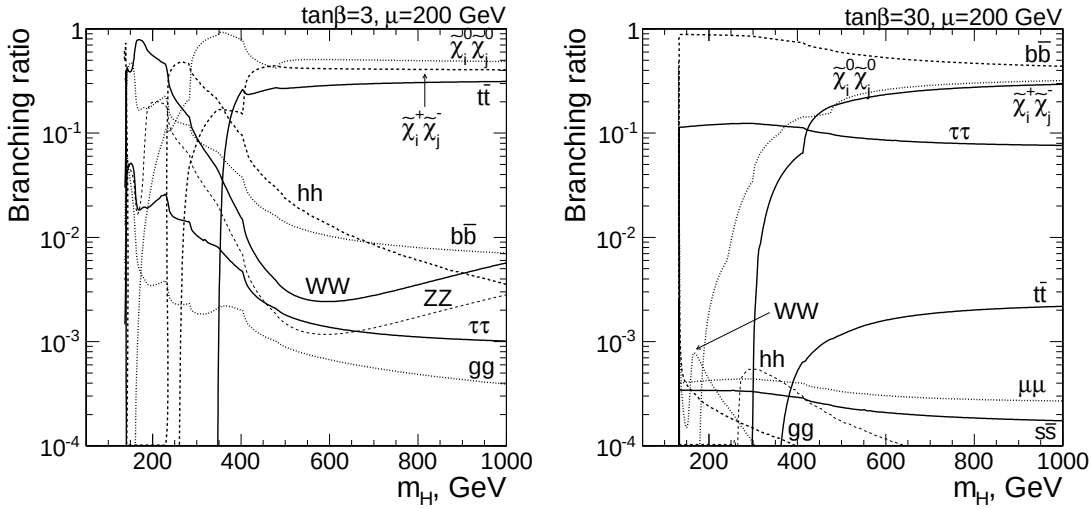

Figure 2.8: Branching ratios of the MSSM heavy scalar Higgs boson H as a function of $\mathrm{m_H}$ for $\tan\beta = 3$ (left plot) and $\tan\beta = 30$ (right plot).

final state albeit the small branching fraction.

The branching ratios of the pseudo-scalar Higgs boson, A, shown in Fig. 2.9 for $\tan\beta = 3$ and 30 and $\mu = 200$ GeV/c^2, behave similarly as those of H, except that the decay modes to vector boson pairs are forbidden. Instead, the decay mode hZ is relatively high for low $\mathrm{m_A}$ and low $\tan\beta$ values. Otherwise, the $\tilde{\chi}_i^0\tilde{\chi}_j^0$, $\tilde{\chi}_i^+\tilde{\chi}_j^-$, $\mathrm{t\bar{t}}$ and $\mathrm{b\bar{b}}$ decay modes are the dominating A decay modes for low $\tan\beta$ values. The branching ratio of the $\tau\tau$ channel is small as for the $\mathrm{H} \rightarrow \tau\tau$ branching ratio for low $\tan\beta$ values.

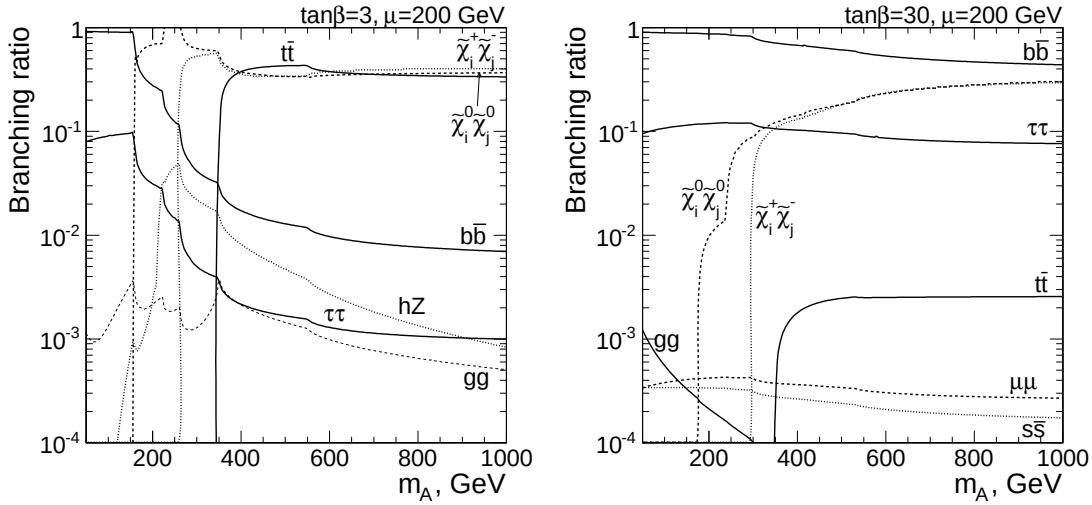

Figure 2.9: Branching ratios of the MSSM heavy pseudo-scalar Higgs boson A as a function of $\mathrm{m_A}$ for $\tan\beta = 3$ (left plot) and $\tan\beta = 30$ (right plot).

The A couplings involving d-type quarks as well as electrons, muons, and τ leptons are directly proportional to $\tan\beta$. Furthermore, the A couplings involving u-type quarks are inversely proportional to $\tan\beta$. Hence, the $\mathrm{b\bar{b}}$ decay mode becomes the most dominating decay mode for A decays at high $\tan\beta$ values. The decay modes containing neutralinos and charginos are the subleading decay modes.

However, like for the H decays, the experimentally most promising decay mode at high $\tan\beta$ values is the $\mathrm{A} \rightarrow \tau\tau$ decay channel which has a substantial branching ratio over the whole $\mathrm{m_A}$ mass range.

The branching ratios of the charged Higgs boson are shown in Fig. 2.10 for $\tan\beta$ = 3 and 30 and μ = 200 GeV/c^2. For low $\tan\beta$ values and $\mathrm{m_{H^\pm}} < \mathrm{m_t}$, the dominating $\mathrm{H^-}$ decay mode is $\tau^-\bar{\nu}_\tau$, which becomes suppressed by the $\bar{\mathrm{t}}\mathrm{b}$ and $\tilde{\chi}_i^0\tilde{\chi}_j^-$ decay modes for $\mathrm{m_{H^\pm}} \gtrsim \mathrm{m_t}$. Since the $\mathrm{H^-}$ coupling to $\tau^-\bar{\nu}_\tau$ depends linearly on $\tan\beta$ and since the $\mathrm{H^-}$ coupling to $\bar{\mathrm{t}}\mathrm{b}$ is proportional to $\mathrm{m_t}\cot\beta$, the branching fraction of the $\tau\nu_\tau$ decay channel is enhanced significantly for high $\tan\beta$ values. For $\mathrm{m_{H^\pm}} \gtrsim \mathrm{m_t}$, the $\bar{\mathrm{t}}\mathrm{b}$ decay mode dominates the $\mathrm{H^-}$ decays for high $\tan\beta$ values with the $\tau^-\bar{\nu}_\tau$ decay mode as the subleading decay mode. Therefore, since the $\bar{\mathrm{t}}\mathrm{b}$ decay mode is difficult to separate from the large QCD multi-jet background, the experimentally most promising decay channel to discover the charged Higgs bosons is the $\tau^\pm\nu_\tau$ decay mode for both low and high $\tan\beta$ values.

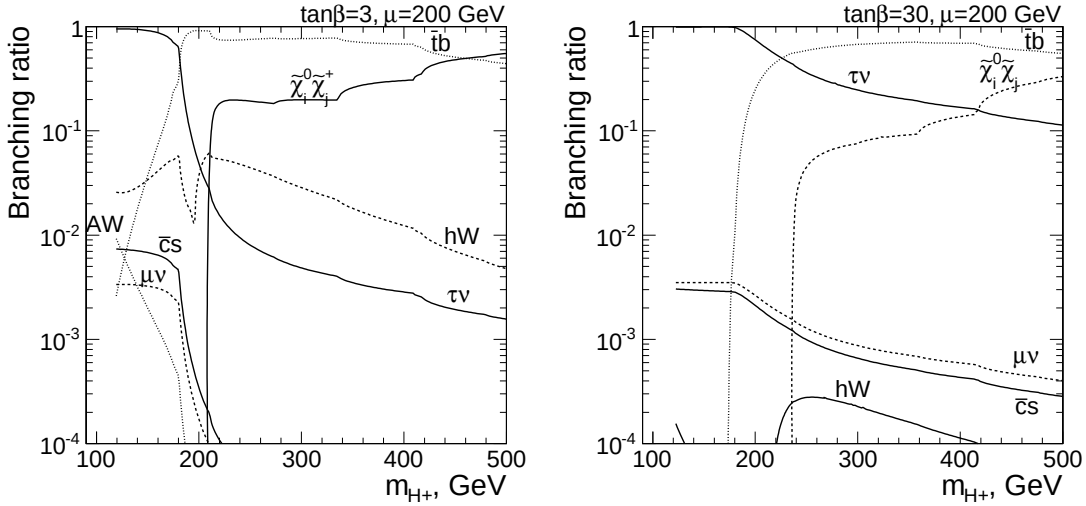

Figure 2.10: Branching ratios of the MSSM charged Higgs bosons $\mathrm{H^\pm}$ as a function of $\mathrm{m_{H^\pm}}$ for $\tan\beta$ = 3 (left plot) and $\tan\beta$ = 30 (right plot).

It should be noted that the branching ratio of the MSSM Higgs bosons to the supersymmetric particles depends decisively on their masses as well as the μ parameter. Since the masses of the lightest neutralino and chargino have been set considerably lower than the TeV energy scale in the $\mathrm{m_h^{max}}$ scenario, the branching ratio to neutralinos and charginos is high over the entire $\mathrm{m_H}/\mathrm{m_A}/\mathrm{m_{H^\pm}}$, $\tan\beta$ parameter spaces. Furthermore, the masses of the squarks and sleptons have been set too high in the $\mathrm{m_h^{max}}$ scenario to produce them in pairs in the Higgs boson decays.

2.2.6 Outlook of the MSSM

Supersymmetrical theories including the MSSM do have some known limitations. They do not explain the soft SUSY-breaking terms or give a mechanism to account for the breaking of the supersymmetry. Furthermore, the supersymmetrical theories including MSSM do not include gravity, and hence they lack the ability to describe all the known particles and interactions between them.

If mixing is allowed between the particle generations and if complex phases are allowed, the soft SUSY-breaking terms yield $\mathcal{O}(100)$ unknown parameters in addition to the standard model parameters. If these parameters are restricted, the number of free parameters is reduced to $\mathcal{O}(20)$ in addition to the standard model ones. Hence, supersymmetrical theories are not at all more elegant than the standard model they seek to extend. However, from a pragmatic point of view, such inelegance might not matter if the theory provides an otherwise accurate description of nature.

To understand the MSSM and supersymmetrical theories, it is necessary to first understand the electroweak symmetry breaking mechanism in them. If the Higgs mechanism does account for the breaking of the electroweak symmetry, it should manifest itself in the minimal form as five Higgs bosons. To discover experimentally these Higgs bosons, the Higgs boson decay modes including τ leptons are essential, if not the most essential, decay modes for all of the five Higgs bosons in a large part of the allowed MSSM parameter space.

2.3 Properties of τ leptons

The potential discovery of the MSSM Higgs bosons via the $\mathrm{H}, \mathrm{A} \rightarrow \tau\tau$ and $\mathrm{H}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decay channels relies on the properties of the τ leptons and their hadronic final states. Identifying and reconstructing hadronic τ decays, i.e. τ jets, is challenging. Firstly, since the τ leptons are usually produced at the interaction point and since they decay very rapidly ($c\tau_{\tau}$ = 87.11 $\pm$ 0.3 μm), their short existence has to be inferred from their daughter particles produced in their decay. Secondly, the τ leptons have several possible decay modes, whose detection requires all the dedicated detectors available in a modern high-energy physics experiment. Thirdly, the τ jets coming from the $\mathrm{H}, \mathrm{A} \rightarrow \tau\tau$ and $\mathrm{H}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decays have to be identified from a very high background of hadronic jets produced in the p-p collisions at the LHC collider and from genuine τ's coming from the $\mathrm{Z} \rightarrow \tau\tau$ and $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decays.

The decay modes of the τ leptons are described in Section 2.3.1. The helicity correlations of the τ leptons, which are important in the $\mathrm{H}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decay channel, are presented in Section 2.3.2.

2.3.1 Branching ratios and final states of τ leptons

The final states and the corresponding branching ratios of the τ lepton decays are shown in Table 2.6. The τ lepton identification strategy is divided into three types of analyses because of the different types of daughter particles coming from the τ decay: the electronic ($\tau \rightarrow \mathrm{e}^{\pm}\nu_{\mathrm{e}}\nu_{\tau}$), the muonic ($\tau \rightarrow \mu^{\pm}\nu_{\mu}\nu_{\tau}$), and the hadronic ($\tau \rightarrow \mathrm{hadr.}$) final states. This thesis concentrates on the hadronic τ decays, i.e. τ jets, which constitute about $\sim$64% of all τ decays.

The hadronic τ decays, which are also called τ jets because of their resemblance with hadronic jets, decay through a chain of vector or axial vector mesons to charged and neutral mesons or directly to one

charged meson. Most τ jets contain either one or three charged particles (one or three prongs). The charged particles are mostly charged pions or with a small fraction charged kaons. As a result of the boost effect, the mesons travel well beyond the dimensions of modern high-energy physics experiments before decaying. Hence, they have to be treated as stable particles albeit their life-times of $\sim 10^{-8}$ s. Final states of five or more charged particles also exist, but their total branching ratio is only $\sim$0.1% of all τ decays. Depending on the decay chain of the τ lepton, a number of π^0's can be produced in the decay. Additionally, a small fraction of τ decays contains K^0's, of which the majority decay after the innermost detector parts because of the boost effect. For the lepton flavor to be conserved, each τ decay involves the production of the elusive τ neutrino, which manifests itself as missing energy when the energy and the momentum conservation are applied to the energy deposition measurement of the particles.

Single τ decay final state	Γ_i/Γ_{tot} (%)	Total (%)
leptonic modes		35.9
$\tau^- \to e^- \bar{\nu}_e \nu_\tau$	17.9	
$\tau^- \to \mu^- \bar{\nu}_\mu \nu_\tau$	17.4	
hadronic modes		
one-prong (excl. K^0's)		48.1
$\tau^- \to h^- \nu_\tau$	11.6	
$\tau^- \to \rho^- \nu_\tau \to h^- \pi^0 \nu_\tau$	26.0	
$\tau^- \to a_1^- \nu_\tau \to h^- \pi^0 \pi^0 \nu_\tau$	9.3	
$\tau^- \to h^- \nu_\tau + \geq 3\pi^0$	1.3	
three-prong (excl. K^0's)		14.6
$\tau^- \to a_1^- \nu_\tau \to 2h^- h^+ \nu_\tau$	9.7	
$\tau^- \to 2h^- h^+ \nu_\tau + \geq 1\pi^0$	5.2	
five-prong (excl. K^0's)		0.1
$\tau^- \to 3h^- 2h^+ \nu_\tau + \geq 0\pi^0$	0.1	
hadronic modes with K^0's		2.0
$\tau^- \to K_S^0 + X$	0.9	
$\tau^- \to K_L^0 + X$	1.1	

Table 2.6: Final states and branching ratios of single τ lepton decays calculated from Ref. [43]. The uncertainty of the branching ratios is 0.1 percent units or smaller.

The final states of di-τ production (Table 2.7) are even more diverse than single τ production final states. The di-τ decays are categorized as leptonic, semi-leptonic and hadronic depending on the decay type of each of the τ leptons. The hadronic di-τ decay mode, which is discussed for the $H, A \to \tau\tau$ channel, constitutes $\sim$42 % of all the di-τ decay modes.

2.3.2 Helicity correlations in the $W^\pm/H^\pm \to \tau^\pm \nu_\tau$ decays

The polarization states of the τ's from the $H^\pm \to \tau^\pm \nu_\tau$ and $W^\pm \to \tau^\pm \nu_\tau$ decays are opposite because of the spin-parity properties of the decaying particle. Exploiting these helicity correlations provides an

Double τ decay final state	$\Gamma_{\mathrm{i}}/\Gamma_{\mathrm{tot}}$ (%)	Total (%)
leptonic modes		12.4
$\tau\tau \to \mathrm{e} + \mathrm{e} +$ neutrals	3.2	
$\tau\tau \to \mathrm{e} + \mu +$ neutrals	6.2	
$\tau\tau \to \mu + \mu +$ neutrals	3.0	
semileptonic modes		45.6
$\tau\tau \to \mathrm{e}$ + hadr. 1-prong	17.9	
$\tau\tau \to \mu$ + hadr. 1-prong	17.4	
$\tau\tau \to \mathrm{e}/\mu$ + hadr. 3-prong	10.3	
$\tau\tau \to \mathrm{e}/\mu$ + hadr. 5-prong	$<$0.1	
hadronic modes (excl. K^0's)		39.6
$\tau\tau \to$ hadr. 1-prong + hadr. 1-prong (without π^0's)	1.3	
$\tau\tau \to$ hadr. 1-prong + hadr. 1-prong (with π^0's)	21.8	
$\tau\tau \to$ hadr. 1-prong + hadr. 3-prong	14.2	
$\tau\tau \to$ hadr. 1-prong + hadr. 5-prong	0.1	
$\tau\tau \to$ hadr. 3-prong + hadr. 3-prong	2.1	
hadronic modes (with K^0's)		2.6
$\tau\tau \to$ hadr. 1/3-prong + hadr. 1/3-prong (with K^0's)	2.6	

Table 2.7: Final states and branching ratios of double τ lepton decays calculated from Ref. [43]. The uncertainty of the branching ratios is 0.1 percent units or smaller.

interesting opportunity to separate τ leptons coming from the $\mathrm{H}^{\pm} \to \tau^{\pm}\nu_{\tau}$ decay from those coming from the $\mathrm{W}^{\pm} \to \tau^{\pm}\nu_{\tau}$ decay. In the following, the helicity correlations are presented separately for the one- and three-prong final states of τ decays. A more detailed description of the τ helicity correlations can be found in Refs. [44, 45, 46].

2.3.2.1 Helicity correlations in one-prong τ decays

It can be seen from Table 2.6, that the dominating hadronic τ decay modes in the one-prong final state are $\tau^{\pm} \to \pi^{\pm}\nu_{\tau}$, $\tau^{\pm} \to \rho^{\pm}\nu_{\tau} \to \pi^{\pm}\pi^0\nu_{\tau}$, and $\tau^{\pm} \to \mathrm{a}_1^{\pm}\nu_{\tau} \to \pi^{\pm}\pi^0\pi^0\nu_{\tau}$, whose branching ratios of all hadronic one-prong τ decays are 24.1%, 54.1%, and 19.3%, respectively, when the decay modes with K^0's are ignored. Decay modes containing more than two π^0's can be considered negligible. Since the kinematics is different for each of these decay modes, they are discussed separately in the following.

Since the charged Higgs boson is a scalar particle and since the neutrinos are left-handed, the scalar object flips the τ helicity to be right-handed in the $\mathrm{H}^{\pm} \to \tau^{\pm}\nu_{\tau}$ decay. On the other hand, in the $\mathrm{W}^{\pm} \to \tau^{\pm}\nu_{\tau}$ decay the gauge interaction holds the helicity conservation and the τ has to be left-handed. Hence, for the $\tau^{\pm} \to \pi^{\pm}\nu_{\tau}$ decay, the pion is produced along the charged Higgs boson direction, whereas in the $\mathrm{W}^{\pm} \to \tau^{\pm}\nu_{\tau}$ decay, the pion is produced in the opposite direction as the W. The differential decay

width of the $\tau^{\pm} \rightarrow \pi^{\pm}\nu_{\tau}$ decay is

$$\frac{1}{\Gamma_{\pi}}\frac{\mathrm{d}\Gamma_{\pi}}{\mathrm{d}\cos\theta} = \frac{1}{2}\left(1 + \mathrm{P}_{\tau}\cos\theta\right), \tag{2.72}$$

where P_{τ} is +1 for τ's from the $\mathrm{H}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decay and -1 for τ's from $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decay and θ is the angle between the pion direction and τ spin quantization axis in the τ rest frame [44]. The τ spin quantization axis is chosen to correspond to the τ momentum direction in the laboratory frame. From equation (2.72) it is easy to see, that the signal τ's are on the average harder than the τ's from the $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decay in the $\tau^{\pm} \rightarrow \pi^{\pm}\nu_{\tau}$ decay mode.

For the τ decay final states including neutral pions, the situation is more complex, since the mesons $\rho^{\pm}$ and $\mathrm{a}_1^{\pm}$ can have both longitudinal and transverse polarization modes. Their differential decay widths are of the form

$$\frac{1}{\Gamma_{\mathrm{v}}}\frac{\mathrm{d}\Gamma_{\mathrm{vL}}}{\mathrm{d}\cos\theta} = \frac{\frac{1}{2}\mathrm{m}_{\tau}^2}{\mathrm{m}_{\tau}^2 + 2\mathrm{m}_{\mathrm{v}}^2}\left(1 + \mathrm{P}_{\tau}\cos\theta\right) \tag{2.73}$$

$$\frac{1}{\Gamma_{\mathrm{v}}}\frac{\mathrm{d}\Gamma_{\mathrm{vT}}}{\mathrm{d}\cos\theta} = \frac{\mathrm{m}_{\mathrm{v}}^2}{\mathrm{m}_{\tau}^2 + 2\mathrm{m}_{\mathrm{v}}^2}\left(1 - \mathrm{P}_{\tau}\cos\theta\right), \tag{2.74}$$

where θ is the angle between the meson direction and the τ line of flight in the τ rest frame, v refers to the vector or axial vector meson in question, and L and T denote the longitudinal and transverse polarization states, respectively. The angle θ, which appears in equations (2.72)- (2.74), can be expressed as

$$\cos\theta = \frac{2\mathrm{X} - 1 - \mathrm{m}_{\pi,\mathrm{v}}^2/\mathrm{m}_{\tau}^2}{1 - \mathrm{m}_{\pi,\mathrm{v}}^2/\mathrm{m}_{\tau}^2}, \tag{2.75}$$

where $\mathrm{X} = \mathrm{p}_{\pi^{\pm}}/\mathrm{E}_{\text{visible } \tau \text{ jet}}$ is the fraction of the visible τ-jet energy carried by the charged pion in the τ rest frame. In the laboratory frame, which is generally used, the variable X is denoted by R_{τ}. From equation (2.75) it is possible to see, that also $\cos\theta$ could be used as a variable to cut on, but the R_{τ} variable is used instead in this study in order to conform with Refs. [44, 45].

In the longitudinal polarization state of the vector or axial vector meson, the meson decay favors uneven distribution of the momentum between the charged and neutral pions. Hence, in the $\tau^{\pm} \rightarrow \rho_{\mathrm{L}}^{\pm}\nu_{\tau} \rightarrow \pi^{\pm}\pi^{0}\nu_{\tau}$ decay, the charged pion is emitted either in the direction of the meson or opposite to it and carries either most or very little of the meson energy. On the other hand, in the transverse polarization state, an even sharing of the momentum of the meson is favored between the pions. Therefore, the charged pion carries in most cases half of the meson momentum in the $\tau^{\pm} \rightarrow \rho_{\mathrm{T}}^{\pm}\nu_{\tau} \rightarrow \pi^{\pm}\pi^{0}\nu_{\tau}$ decay. The $\mathrm{a}_{1\mathrm{L,T}}^{\pm}$ decays follow the same principle as the $\rho_{\mathrm{L,T}}^{\pm}$ decays, except that the presence of the second neutral pion in the final state produces more smeared R_{τ} distributions. In the $\tau^{\pm} \rightarrow \mathrm{a}_{1\mathrm{L}}^{\pm}\nu_{\tau} \rightarrow \pi^{\pm}\pi^{0}\pi^{0}\nu_{\tau}$ decay, the charged pion carries in most cases very little of the meson energy, but has also some chance to carry most of the meson energy. On the other hand, on average one third of the meson momentum is carried by the charged pion in the $\tau^{\pm} \rightarrow a_{1\mathrm{T}}^{\pm}\nu_{\tau} \rightarrow \pi^{\pm}\pi^{0}\pi^{0}\nu_{\tau}$ decay. The small contributions from K^{*} and $\mathrm{K}^{\pm}$ in the the τ decays lead to similar effects. The differential decay widths of the $\rho_{\mathrm{L,T}}^{\pm}$ and $\mathrm{a}_{1\mathrm{L,T}}^{\pm}$

decays are shown in Figure 2.11 in the τ rest frame. The $\tau^{\pm} \rightarrow \pi^{\pm}\nu_{\tau}$ decay mode can be added to the figure as a δ-function at $\mathrm{R}_{\tau} \sim 1$. The distributions are similar in the laboratory frame.

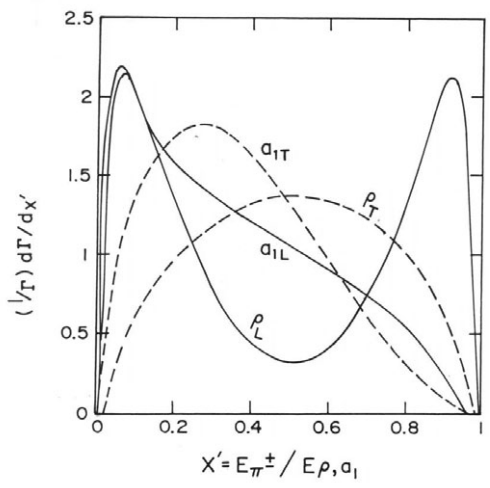

Figure 2.11: Distributions of the $\rho^{\pm}_{\mathrm{R}} \rightarrow \pi^{\pm}\pi^{0}$ and $\mathrm{a}^{\pm}_{1\mathrm{R}} \rightarrow \pi^{\pm}\pi^{0}\pi^{0}$ decay widths in the energy fraction carried by the charged pion, shown separately for the transverse and longitudinal states of $\rho^{\pm}$ and $\mathrm{a}^{\pm}_{1}$ polarization. Taken from Ref. [45].

From the equations (2.73)-(2.75) it can be seen, that requiring the jet $\mathrm{E_T}$ to exceed a threshold biases the fraction of longitudinally and transversally polarized vector or axial vector mesons. A high threshold on τ jet $\mathrm{E_T}$ value enhances the longitudinal polarization states of the mesons in the $\mathrm{H}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decays ($\mathrm{P}_{\tau} = 1$, $\cos\theta \rightarrow 1$) and the transverse polarization states for the $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decay ($\mathrm{P}_{\tau} = -1$, $\cos\theta \rightarrow 1$). Hence, the τ's from $\mathrm{H}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decays can be separated from the $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decays by requiring the charged pion to carry either very little or most of the visible τ jet energy. In practice, however, much of the events in which the charged pion carries very little of the visible τ jet energy are lost as a result of a cut on the leading track $\mathrm{p_T}$.

It has also been suggested in Ref. [44], that if the neutral pions could be efficiently reconstructed, it would be possible to construct a discriminating variable based on the even/uneven momentum distribution between the pions such as $\Delta\mathrm{p_T} = |\mathrm{p}^{\pi^{\pm}}_{\mathrm{T}} - \mathrm{p}^{\pi^{0}}_{\mathrm{T}}|$. The momentum difference is expected to be large for τ's from $\mathrm{H}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decays and small for τ's from the $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decay because of the uneven (even) vector or axial vector meson momentum sharing between the pions in the transverse (longitudinal) polarization state of the meson.

2.3.2.2 Helicity correlations in three-prong τ decays

From Table 2.6 it can be seen, that the two most important three-prong τ jet final states are the $\tau^{\pm} \rightarrow \mathrm{a}^{\pm}_{1}\nu_{\tau} \rightarrow \pi^{\pm}\pi^{\pm}\pi^{\mp}\nu_{\tau}$ and $\tau^{\pm} \rightarrow \pi^{\pm}\pi^{\pm}\pi^{\mp}\nu_{\tau} + n \times \pi^{0}$, whose branching ratios are 66.4% and 35.6% of all three-prong three-prong τ jets, respectively, when K^{0}'s are omitted. In the following, the

former decay mode is considered, since it is the predominant one.

The helicity correlations for the $\tau^{\pm} \rightarrow \mathrm{a}_1^{\pm} \rightarrow \pi^{\pm}\pi^{\pm}\pi^{\mp}$ decays follow the helicity correlations of the $\tau \rightarrow \mathrm{a}_1\nu_\tau$ decay mode described in Section 2.3.2.1 for the one-prong τ decays with the exception, that now all three pions have a charge. If a high enough jet $\mathrm{E_T}$ threshold is requested, the τ's from the $\mathrm{H}^{\pm} \rightarrow \tau^{\pm}\nu_\tau$ decay will be predominantly longitudinally polarized and the τ's from the $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_\tau$ decay predominantly transversally polarized. The longitudinally polarized a_1 decays favor uneven energy sharing between the pions, whereas the transversally polarized a_1 decays favor even energy sharing between the pions. This difference can be exploited to construct a variable to separate differently polarized τ's.

One approach is to take advantage of the fact, that in $\mathrm{H}^{\pm} \rightarrow \tau^{\pm}\nu_\tau$ decays, the two same signed pions carry either the most or very little of the visible τ jet energy. On the other hand, the two same signed pions carry $\sim \frac{2}{3}$ of the visible τ jet energy for τ's coming from the $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_\tau$ decay, because of the even energy sharing between the pions. The fraction of energy carried by the same signed pions can be defined as $\mathrm{R}_{\tau 3} = \mathrm{p}_{\pi^{\pm}\pi^{\pm}}/\mathrm{E}_{\mathrm{visible}\ \tau\ \mathrm{jet}}$ [46].

3 CMS - the Compact Muon Solenoid experiment at the LHC

The Compact Muon Solenoid (CMS) [47, 48] shown in Fig. 3.1 is one of the largest detectors built for the LHC. The detector has a diameter of 15 m, an overall length of 22 m, and a total weight of 14500 tons. Its name is derived from the chosen magnetic field design of a superconducting solenoid coil, which requires less space than the choice of a toroidal magnetic field. The main design goals of CMS are a very good muon system, the best possible electromagnetic calorimeter, and a high-quality central tracking system, as well as a financially affordable detector. The CMS is designed to study the breaking of the electroweak symmetry by searching for the Higgs boson and supersymmetric particles. The search for CP violation and the study of quark-gluon plasma with heavy ions are also included in its physics objectives. Additionally, the CMS is capable of diffractive physics together with the TOTEM experiment [49]. Hence, the CMS experiment is an adaptable universal-purpose detector.

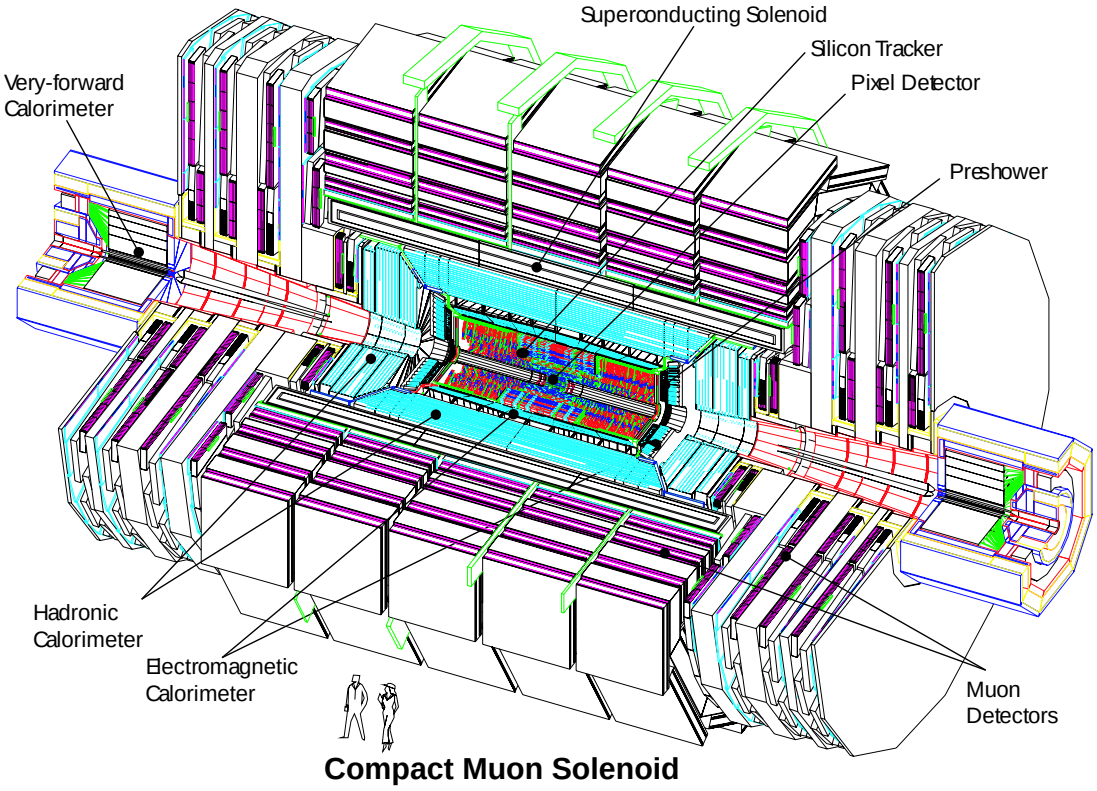

Figure 3.1: A perspective drawing of the CMS experiment with labels for the subdetectors.

The geometry of CMS follows that of a standard barrel-shaped collider experiment setup. The proton or heavy ion beams are guided to cross each other at the center of the detector and the surrounding subdetectors record all feasible information from the debris produced in the collisions. The solenoid magnet is aligned along with the beam line and the subdetector modules are located concentrically around the beam line both in- and outside the solenoid structure. The concentric layers of detectors

are complemented at both ends of the barrel with endcap discs that close the barrel structure. Also the endcap discs are equipped with detectors in order to provide nearly hermetic measurements of the collisions.

The innermost structure of the CMS detector is a thin beam pipe, which houses the ultra high vacuum of the LHC accelerator ring and inside which the particle bunches cross each other. The beam pipe is surrounded by the tracker which consists of silicon pixels and silicon microstrip detectors to measure the tracks of the charged particles produced in the collisions. The tracker is surrounded by the calorimetry system. The innermost component of the calorimetry system is the electromagnetic calorimeter (ECAL), which provides energy measurements of electrons, positrons, and photons. The outermost component of the calorimetry system is the hadronic calorimeter (HCAL), which measures the energy of particles with enough energy to pass through the electromagnetic calorimeter.

The tracker and the calorimetry system are compact enough to fit inside the superconducting coil of the solenoid magnet which is designed to produce a homogenic magnetic field of 4 T transverse to the beam line. The high magnetic field is required to distinguish the electric charge of the particles and to measure their momentum from the curvature of their tracks. Additionally, the magnetic field causes particles with a small transverse momentum to spiral before reaching the first tracking detectors. Outside the solenoid, the high magnetic field is relayed back through a massive iron yoke which surrounds the solenoid in three layers both in the barrel and the endcap regions. The iron yoke also acts as an absorber for any particles escaping the detector.

The outermost detectors, the muon stations, are located between and outside the layers of the iron yoke. They identify the muons, which are the only charged particles passing through the calorimeters and the iron yoke. The information provided by the muon stations is combined with the tracker information to obtain excellent muon momentum resolution.

The subdetectors must also be fast enough to read the detector signals within 25 ns, before the next beam particle bunches cross each other. Since the total cross-section for proton-proton collisions at the designed center of mass energy of 14 TeV of the LHC is of the order of 100 mb, and since the design luminosity of the LHC is 10^{34} cm^{-2}s^{-1}, the detectors will have to be able to observe of the order of 10^9 inelastic events per second. Of these events, an average of about 20 events occurs during the same beam bunch crossing. Such a high rate of events calls for special solutions for selecting and storing the interesting events. A sophisticated, flexible, reliable, and fast triggering system is required to drop the stored event rate from 40 MHz to $\sim$100 Hz, which is a feasible rate for recording the events to storage media. The rate of stored events still poses a challenge for computing and storing techniques, because the yearly data volume which has to be accessible from practically all around the world will amount to about one petabyte (10^{15} bytes). Furthermore, the high rate of events causes the subdetectors to be constantly irradiated by a high flux of particles. This irradiation requires the selection of special radiation hard detectors.

In the following, the tracker, calorimetry system, and muon systems are described in more detail in Sec-

tions 3.1, 3.2, and 3.3, respectively. Furthermore, the triggering and data-acquisition are presented in Section 3.4.

3.1 Tracking system

The innermost part of the CMS detector is the tracker [47, 50, 51] which has a length of 5.8 m and a radius of 1.3 m. The task of the tracker is to measure the tracks of charged particles produced in the particle collisions with high precision. The particles with enough momentum in the plane transverse to the beam axis emerge through the thin beryllium beam pipe which houses the vacuum of the LHC collider and propagate through the tracker modules before reaching the calorimetry system. The tracker plays thus an important role in the measurement of the particle momenta and charge, as well as in the reconstruction of the decay vertices of short lived particles. These properties are essential for the identification of both short- and long-lived particles.

Since most particles which are produced in the hard interaction of the beam particles have a very short mean life time even with the boost effect, the tracker is required to have a very good position resolution in order to reconstruct the production and decay vertices of these elusive particles. Additionally, the tracker has to be constructed from as little material as possible to minimize multiple scattering, bremsstrahlung, photon conversions, and nuclear interactions. The high interaction rate at the LHC also requires the tracker detectors to withstand radiation damage, since they will absorb a dose equivalent to 1.6×10^{14} neutrons/cm^2 (eq. 1 MeV) during the estimated operation time of ten years.

To fulfill the above criteria, as well as readout speed and granularity, the tracker has been built of silicon micro-strip and pixel sensors. The layout of these sensors in the tracker is shown in Fig. 3.2. To facilitate cabling and cooling services, the tracker space has been divided to barrel ($|\eta| < 1.48$) and endcap regions ($1.48 < |\eta| < 2.50$). At the heart of the tracker is the pixel detector, which has a small enough granularity to cope with the high occupancy environment. The pixel detector is surrounded with silicon strip sensors. The barrel region of the tracker is divided into the tracker inner barrel (TIB) and tracker outer barrel (TOB) parts in which the sensors have been place on concentric layers around the beam axis, as seen in Fig. 3.2. The endcap regions are divided similarly to tracker inner discs (TID) and tracker endcaps (TEC), which are placed perpendicularly towards the beam line. The placement of the barrel and endcap sensors has been designed so that each particle will pass through at least ten sensors independent of the pseudo-rapidity supposing the particles are long-lived enough and have a high enough transverse momentum.

Overall, the CMS tracker is equipped with $\sim$15000 silicon strip sensors, which cover an area of 206 m^2, making it the largest silicon based tracker built so far. In the following, the most important aspects of the silicon pixel and silicon strip sensors are described. More details on the engineering and performance of the tracker can be found in Refs. [47, 52].

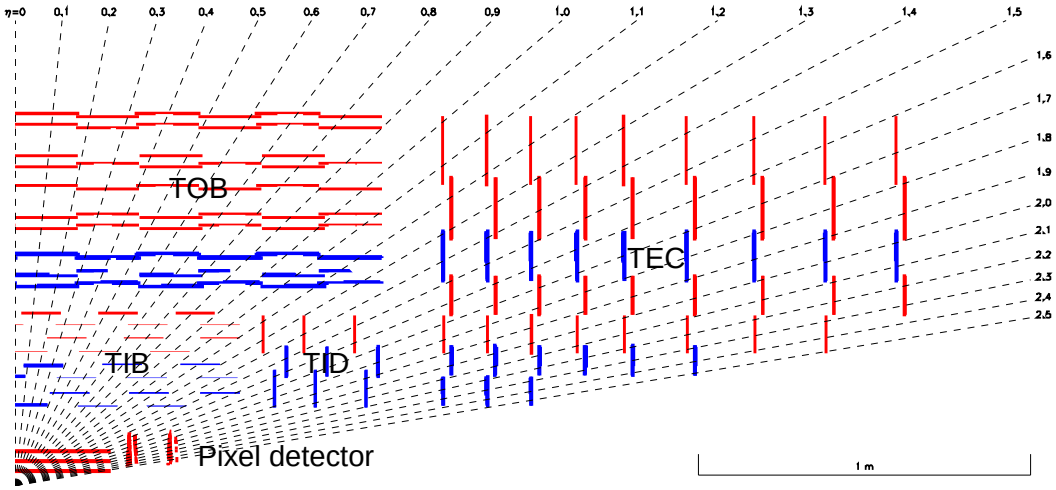

Figure 3.2: Schematic cross-section of one quarter of the CMS tracker in the r,z. The red and blue lines correspond to single and double-sided sensor modules, respectively.

3.1.1 Silicon pixel sensors

The pixel detector [53, 54], which is the innermost part of the tracker, consists of quadratical arrays of n+ implants on n- substrates with a thickness of 250 μm. As a charged particle propagates through the depletion region of the junction, it produces electron and hole pairs which are collected to the readout chips with an electric field. The cloud of charge carriers produced by a charged particle passing through the sensor is diffused also to the neighboring pixel cells, which allows a spatial resolution of $\sim$15 μm in both directions.

The area of the individual pixel implants is 100$\times$150 μm^2. The near quadratical shape of the pixel implants has been chosen to achieve the optimal resolution in both r,ϕ and z coordinates. Each pixel implant is bump-bonded to a readout chip, which is glued onto a 270 μm thick silicon base plate attachable to a cooling frame. Each readout chip reads an array of 53$\times$52 pixel rows and columns. The total number of readout channels for the whole pixel detector is about 6.6$\times$10^7. The high granularity of the pixel implants guarantees an average occupancy of about 10^{-4} per pixel implant per LHC bunch crossing.

The pixel detector surrounds the interaction point at mean radii of 4.4, 7.3, and 10.2 cm in the barrel region. The 53 cm long barrel part is complemented on both ends of the barrel with two endcap discs. In the endcap discs, the pixel sensors are mounted on blades, which are rotated 20$^\mathrm{o}$ around their radial axes to benefit from the Lorentz effect, i.e. the diffusion of the charge carriers in the high magnetic field to the neighboring pixel cells. Because of the high dose absorbed at these radii, the pixel sensors must be replaced with new ones during the CMS operation period. The effect of radiation damages and leaking currents has been minimized with different configurations of guard rings around the pixel implants [55].

The pixel detector is essential for the reconstruction of secondary vertices from b hadrons and τ decays. The layout of the pixel detector guarantees each track produced within the interaction region to cross at

least three pixel sensors for $|\eta| < 2.2$ and at least two pixel sensors for $2.2 < |\eta| < 2.5$. Hence, the pixel detector may be used to perform relatively fast track finding which may be used to set coarse constraints for selecting interesting events at the trigger.

3.1.2 Silicon microstrip sensors

The inner and outer tracker layers consist respectively of four and six concentrical layers of silicon microstrip sensors in the barrel region. The barrel layers are complemented with three and seven endcap discs at both ends of the inner and outer barrel layers, respectively. The choice to use silicon microstrip sensors was based on building a compact, efficient, high-resolution, reliable, and a cost-efficient tracker. Silicon microstrip sensors are very fast, they have an excellent resolution, and the latest generation of them is also resistant enough against radiation damages. Additionally, silicon microstrip sensors do not need extra structures such as gas feeding lines, and they may be operated at nearly room temperatures.

Because of the high number of sensors needed, it was necessary to use single-sided p-on-n configuration for the sensors. Such configuration could be manufactured reliably and cost-efficiently on industrial production lines for 6" wafers. The crystal lattice configuration $<100>$ was chosen for the silicon in order to minimize the negative effects of the high radiation operating environment [56]. To minimize leakage current from radiation damage to the crystals, the whole tracker is cooled to a temperature of -20^{o}C.

The sensors of the inner tracker were selected to consist of 320 μm thick substrates with a maximum strip length of 12 cm with a pitch of 80-140 μm, which yields an average occupancy of 3-4% per LHC bunch crossing. The sensors of the outer tracker were designed to have a maximum strip length of 16 cm and a thickness of 500 μm to compensate for the increase of noise caused by the longer strips. The pitch of the sensors is varied between 80-180 μm, which yields the mean occupancy to 1% per LHC bunch crossing. To enhance a two dimensional measurement of the hits, the first and second barrel layers in both the inner and the outer tracker regions, as well as in the endcaps at the equivalent radii, are equipped with stereo sensors. These consist of two single-sided sensor elements coupled back-to-back so that the second sensor is rotated with 100 mrad with respect to the other.

3.2 Calorimetry system

The CMS calorimetry system [47] shown in Fig. 3.3 consists of an electromagnetic calorimeter (ECAL) surrounded by a hadronic calorimeter (HCAL). The task of the calorimety system is to measure the direction and the energy of all the particles produced in the collisions in the region of $|\eta| < 5.2$ except for muons and neutrinos. The hermeticity of the calorimetry system is essential for determining the presence and the direction of neutrinos from the missing energy through the vector sum of energy deposition in the transverse plane.

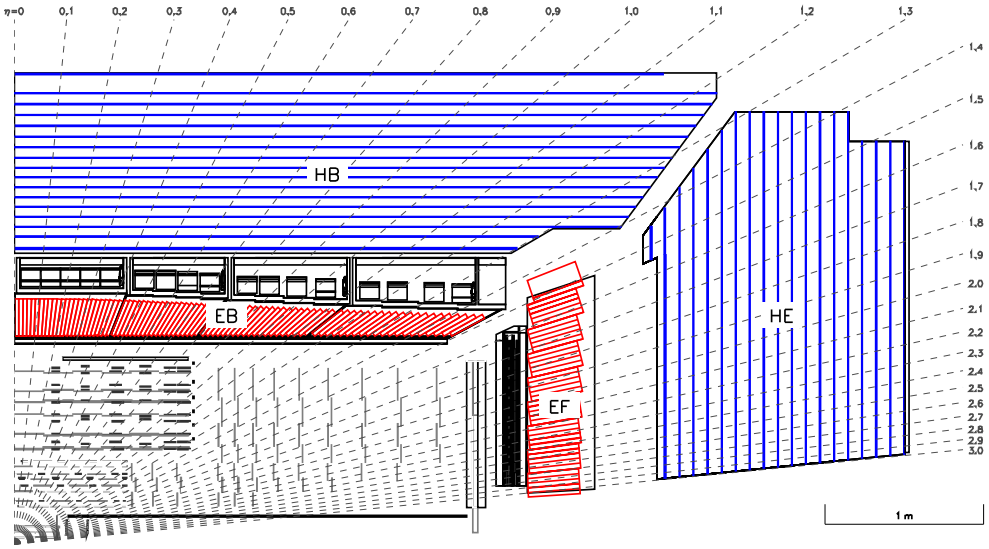

Figure 3.3: A quarter view of the CMS calorimetry system in the r,z plane. The barrel and endcap parts of the ECAL (red color) and HCAL (blue color) are shown. The dashed lines correspond to different pseudo-rapidity values.

The electromagnetic calorimeter is divided into barrel (EB) and endcap (EE) parts which cover the region $|\eta| < 1.5$ and $1.5 < |\eta| < 3.0$, respectively, as seen in Fig. 3.3. In the region $1.7 < |\eta| < 2.6$, the ECAL is preceded by a preshower detector, which provides additional granularity for measuring the position of electrons and photons. The hadronic calorimeter is also divided into barrel (HB) and endcap (HE) parts as seen in Fig. 3.3. These parts are complemented with forward calorimeters, which are located at both ends of the CMS detector and which extend the calorimetry coverage up to $|\eta| = 5.2$. Furthermore, an outer calorimeter, which is located outside the magnet coils within the region $|\eta| < 1.3$, increases the HCAL thickness at small incident angles.

In the following, the ECAL and HCAL are described briefly. More details on the engineering and performance of the calorimetry system can be found in Refs. [47, 52, 57]

3.2.1 Electromagnetic calorimeter

The task of the electromagnetic calorimeter [58] is to measure precisely the energy and position of photons and electrons and also to identify these particles. Its energy measurement relies on the electromagnetic interaction of these particles in the active medium of the calorimeter Additionally, one of the design goals of CMS was to construct the best possible electromagnetic calorimeter for the search for the Higgs boson in the $\mathrm{h} \rightarrow \gamma\gamma$ decay channel, which requires a good separation of single photons from π^0 decays. To fulfill these tasks, the ECAL is constructed from homogeneous lead tungstate ($PbWO_4$) crystals, which have a fine granularity, short radiation length, and a fast readout. Furthermore, the crystals are radiation resistant.

The crystals in the barrel region have a front face of 22×22 mm^2 which matches well the Molière radius of 21.9 mm. This design corresponds to a transverse granularity of $\Delta\eta \times \Delta\phi$ =0.0175×0.0175 in the barrel region increasing progressively in the forward regions up to $\Delta\eta \times \Delta\phi$ =0.05×0.05. The rear face of the crystals is 26×26 mm^2. To maximize the energy deposition within the crystal, the front face of the crystals is aligned perpendicular to the interaction point. The length of the crystals has been chosen to be about 23 cm which corresponds a total thickness of 25.8 radiation lengths. In the endcaps, the front and rear face of the crystals is 29×29 mm^2 and 30×30 mm^2, respectively. The length of the crystals in the endcap regions is 22 cm which corresponds to 24.7 radiation lengths. The photons of the electromagnetic showers are read with avalanche photodiodes and vacuum phototriodes in the barrel and endcap regions, respectively. The crystals emit ∼80% of their stored light within the 25 ns bunch crossing time of the LHC to the photodetectors.

The crystals are contained within a support structure made of glass fiber-epoxy resin. The support structures are grouped as modules which cover 20° in ϕ and which contain 400 or 500 crystals depending on the η position. The modules are separated from each other by 4 mm thick aluminum plates and grouped further to supermodules. In the endcaps, groups of 5×5 crystals are grouped as modules. Although the ECAL has been built to be as hermetic as possible, it is a known effect, that a small amount of electrons may escape through the cracks between the modules and the gap in between the barrel and endcap detectors to the HCAL. The barrel and endcap regions are populated with 61200 and 14600 crystals, respectively. Since the light yield of the crystals depends on their temperature, they are cooled to 18°C to ensure a stable energy resolution.

The preshower detector, which complements the ECAL at $1.65 < |\eta| < 2.6$, adds another three radiation lengths to the ECAL depth. The purpose of the preshower detector is to initiate the electromagnetic cascade with thin lead radiators and to measure the spatial energy distribution with position sensitive silicon sensors. About 95% of the electrons start showering already after the first of two lead radiators of the preshower detector. The extra granularity is necessary to separate π^0 decays from single photons, since the two photons from the π^0 decay are highly collimated in the forward directions.

3.2.2 Hadronic calorimeter

The aim of the hadronic calorimeter [59] is to measure the energy and direction of jets of particles arising from the hadronization of quarks or gluons. The HCAL plays an important role also in the identification of neutrinos or exotic particles by measuring the missing transverse energy. By co-operating with the electromagnetic calorimeter and the muon system, the hadronic calorimeter also helps in the identification of electrons, photons, and muons.

The hadronic barrel and endcap calorimeters are sampling calorimeters. The barrel part consists of 51 and 57 mm thick brass (70% Cu, 30% Zn) absorbers interleaved with the active medium of 4 and 9 mm thick plastic scintillator tiles placed in 17 layers. The innermost and outermost layers are made of 61

and 75 mm thick stainless steel, respectively, to add structural strength. In the endcaps, 79 mm thick brass absorbers are used interleaved with 9 mm thick plastic scintillator tiles. The energy measurement of the HCAL relies on nuclear interactions of the incident particles with the absorbing material. Brass was chosen as the main absorber material because of its density, which yields a relatively short nuclear interaction length, and also because of its non-magnetic properties. The effective thickness of the HCAL is 5.8 interaction lengths at $\eta = 0$ increasing as $1/\sin\theta$, where θ is the azimuthal angle, to 10.6 interaction lengths at $|\eta| = 1.3$. In the endcaps, the effective thickness remains at $\sim$10 interaction lengths. To compensate for the lower effective thickness at small η values, additional scintillation layers are placed outside the superconducting coil. This extra absorber increases the effective thickness of the HCAL to $\sim$11 interactions lengths also for the small η values. The ECAL adds 1.1 interaction lengths to the HCAL effective thickness.

The granularity of the scintillator tiles is $\Delta\eta \times \Delta\phi = 0.09\times0.09$ up to $|\eta| = 1.6$ and increasing to $\Delta\eta \times \Delta\phi = 0.18\times0.18$ in the endcaps. The light yield in the tiles is transported via wavelength shifting fibers outside the wedges and read out with hybrid photodiodes. Based on the dimensions of the scintillator tiles, the whole calorimetry up to $|\eta| = 3.0$ is covered with a grid of 56×72 tiles in η and ϕ directions, respectively.

The forward calorimeters are constructed out of a 1.65 m thick steel absorber, which corresponds to an effective thickness of $\sim$10 radiation lengths. Embedded in the steel absorber, quarz fibers run at a spacing of 5 mm to each other parallel to the beam axis. The energy of the particle jets is measured from the Cherenkov light produced in the quartz fibers. The light produced in the fibers is transported with a wave guide into a shielded location, where the light is read out with photomultiplier tubes. Since half of the fibers run through the whole absorber length and the other half start after 22 cm of absorber, it is possible to separate electromagnetic and hadronic showers. The average deposited energy in the forward calorimeter is $\sim$700 GeV per LHC bunch crossing, which is an unprecedented amount so far in high energy physics.

3.3 Muon system

The task of the muon system [60] is to identify muons which pass through the detector material, to measure their momentum, and to enable the triggering of event based on the presence of muons. The muons are useful, because they provide a clean signature for many interesting physics channels. The rate of cosmic muons, which could fake muons produced in the collisions, is reduced by the $\sim$80 m of solid rock on top of the CMS detector and the remaining cosmic muons can be effectively suppressed by requiring, that the muons originate from the interaction point. Furthermore, the muon system can be used to require muon free regions to suppress certain decay channels. Since the effective depth of the material between the interaction point and the muon system is about 16 interaction lengths, the punchthrough from hadronic particles is negligible.

The muon system is shown in Fig. 3.4. The muon detectors are interleaved in both the barrel and the endcap regions with iron slabs of the yoke which returns the magnetic field to the solenoid. Three different types of muon detectors are used. In the barrel region, the muons are measured with drift tubes (DT) up to $|\eta| = 1.2$, whereas the endcap region is populated with cathode strip chambers (CSC) up to $|\eta| = 2.4$. The drift tubes and cathode strip chambers are complemented with resistive plate chambers (RPC) which provide an independent measurement of the muons. The independent measurement is needed because of uncertainty in the background rates and in the ability of the muon system to measure the correct beam-crossing time when the LHC is operated at full luminosity. The resistive plate chambers cover the barrel and endcap regions up to $|\eta| = 1.6$. The range $1.6 < |\eta| < 2.4$ is planned to be covered with additional RPC's once the full design luminosity of the LHC has been reached.

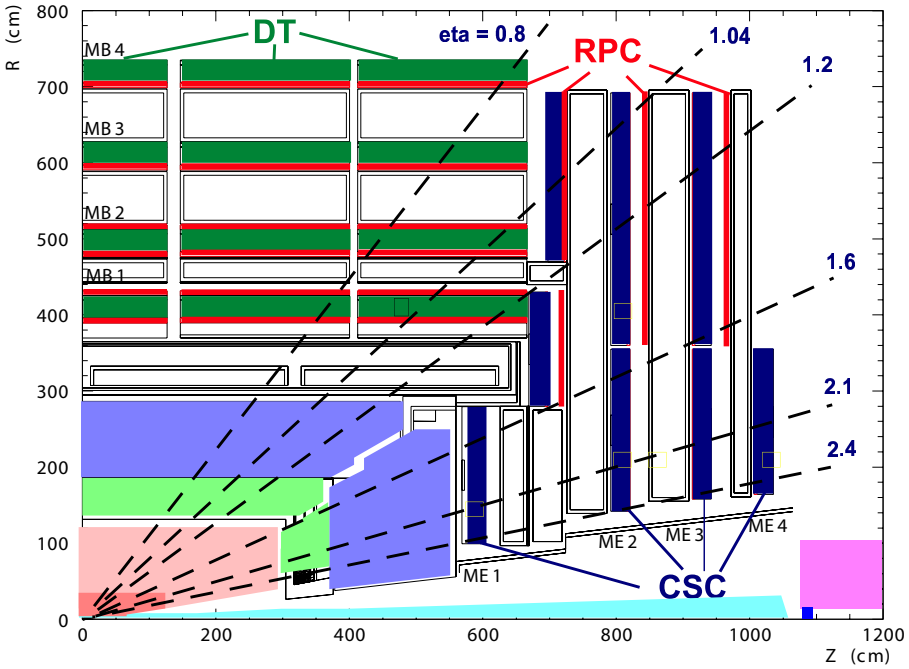

Figure 3.4: Schematic view of the muon system in the r,z plane. The dark green, dark red and dark blue regions correspond to the DT, RPC and CSC muon chambers, respectively.

For each muon, four measurement points are provided by the DT and CSC detectors, except for a gap between the barrel and endcap regions at $|\eta| = 1.4$, where three measurement points are obtained, as can be seen in Fig. 3.4. Additionally, the RPC's yield six measurement points up to $|\eta| = 0.8$ with the measurements decreasing to three, when $|\eta| = 1.6$ is reached. Hence, a relatively good muon momentum estimate is obtained from the muon measurements alone. An even better momentum resolution is obtained, when the muon system measurements are combined with the tracker.

The muon system contains $\sim$25000 m^2 of active detector medium and about one million measurement channels. More details of the engineering and performance of the muon system can be obtained from Refs. [47, 52].

3.4 Triggering and data-acquisition

At the interaction point inside the CMS detector, two proton bunches moving in the opposite directions inside the vacuum tube of the LHC will cross each other every 25 ns, which corresponds to a rate of 40 MHz. At the LHC design luminosity of $\mathcal{L} = 10^{34}$ cm^{-2}s^{-1}, each bunch crossing produces on average $\sim$20 interactions. Most of these interactions are just nearby hits, where the protons scatter diffractively to a very forward direction. Although the diffractive phenomena can be used to measure interesting things such as proton structure, the rare cases of head-on collisions are much more interesting instead of the nearby misses, since more energy is localized in a point-like state and hence available for conversion into new particles. To select only the most interesting events and to reduce the initial event rate of 40 MHz down to $\sim$100 Hz, which is the maximum rate for the online computer farm to store the data, a sophisticated triggering system is needed.

The CMS trigger [61, 62] is divided into hardware and software based triggers known as first level (Level-1) and high level triggers (HLT). Since the measurement signals propagate just a few meters in the cables before the next bunch crossing occurs, the Level-1 trigger collects the signals from the sub-detector system readout electronic into a pipelined buffer. The pipelined buffer gives the Level-1 trigger 3.2 μs time to make a decision with special fast processors whether to examine the event further or to discard it. The decision is based on regional data from the calorimeters and muon chambers. The Level-1 trigger is designed to reduce the event rate by a factor of 4000 down to 100 kHz.

The final event rate reduction to $\sim$100 Hz is achieved by the high level trigger, which uses an onsite online computer farm equipped with commercial processors. The high level trigger employs all available data of the subdetectors and applies special software algorithms to the event to select only the most interesting ones. It is sometimes divided into Level-2, Level-2.5, and Level-3 triggers based on the type of data the algorithm utilizes. The Level-2 algorithms are based on data from the calorimetry and muon systems and feature more detailed reconstruction of physics objects than at the Level-1 trigger. The Level-2.5 and Level-3 triggers are based on partial and full tracker data, respectively. In practice, there is no sharp division between the HLT steps. The decision to keep an event is based on the result of matching the event with predefined analysis data sets based on the physics goals of the CMS detector. The bandwidth allowed for b physics, top physics, Higgs boson searches, search for supersymmetric particles, search for exotica, and data needed for detector calibration have to be agreed upon within the physics groups.

More detailed description of the data-acquisition system and the algorithms used for triggering can be found in Refs. [47, 52, 61, 62]. The stored data is reconstructed and shared with the universities and institutes via the worldwide LHC computing grid (WLCG) [63]. The event data format and the CMS computing system are described in detail in Ref. [64].

4 Simulation and reconstruction of particle collisions

This chapter describes the tools and methods used for simulating high-energy physics collision events in the CMS experiment. The full Monte Carlo (MC) simulation and analysis chain is shown schematically in Fig. 4.1 with a list of the most important software packages. Prior to the experiment being started and experimental data received, the simulation chain provides the only means to acquire valuable knowledge for designing the experiment, its triggering system, determining the most interesting physics channels in the pursuit of new physics, and preparing for the data analysis. Once experimental data from the experiment are available, it is possible to use the same reconstruction and analysis tools developed for handling the simulated data. It should be noted that the simulation chain does not become obsolete once the experiment is running, because simulation data concurrent with experimental data are still important for determining the algorithm uncertainties, measuring the expected backgrounds, and adjusting the triggering parameters. The simulation and experiment are thus interconnected through an iterative loop and the interpretation of the experimental data will always be based to some degree on assumptions made with the simulation models.

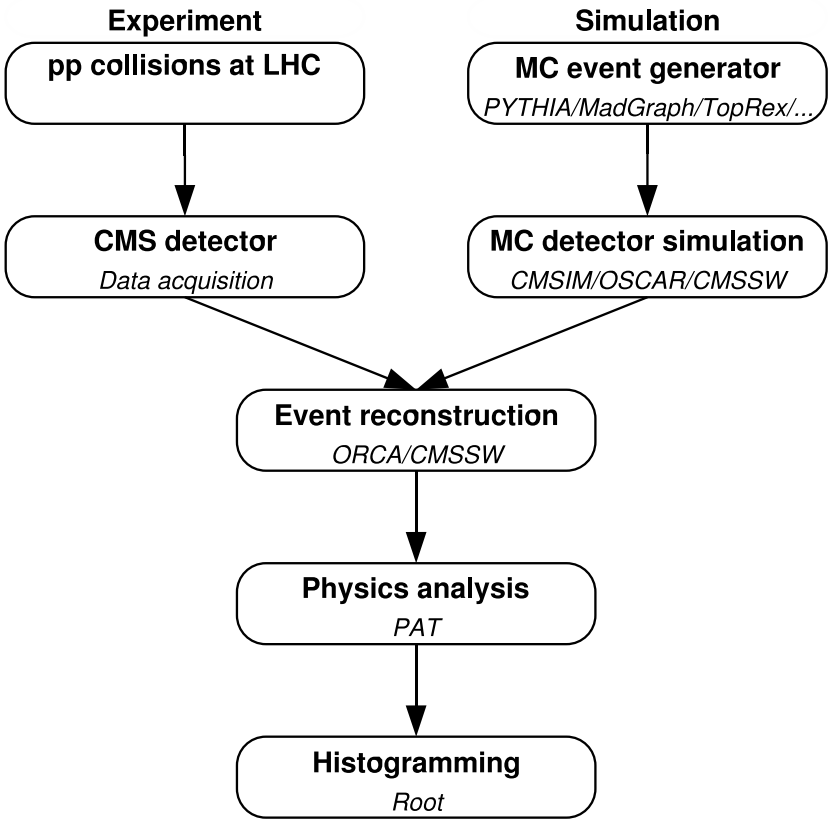

Figure 4.1: Sketch of the MC event simulation and analysis chain.

When the experiment is running, data are collected from the proton-proton collisions by making mea-

surements with the subdetectors and by digitizing the measurements with the corresponding readout electronics. In MC simulation, the events are generated by at least one event generator. The generated events are then resimulated with the GEANT [65] detector description to obtain material effects and to receive the digitized measurement signals. The digitized events are reconstructed in detail with event reconstruction software and analysed with proper analysis software to produce the results of the subject being studied.

Each phase of the simulation requires notable computing resources in order to produce results with high enough statistics. The need for computing resources is minimized with persistent objects which are stored for later analysis. The requirements for notable processing and data storage resources, as well as data transfer capabilities, are met with the use of a data grid [66], a high-speed hierarchical computing network distributed across the world.

In the following, the generation of MC events is described in Section 4.1, the detector simulation is presented in Section 4.2, and the reconstruction and analysis of the events is discussed in Section 4.3. Also the reconstruction of the most important basic physics objects is described in Section 4.3. The reconstruction of higher level physics objects which rely on the basic physics objects will be discussed in later chapters. Finally, the methods for validating the simulated and reconstructed data are discussed in Section 4.4

4.1 Generating events

The purpose of MC event generators is to simulate a given particle reaction according to known or hypothetical physics models. The simulations usually use MC techniques and they are conducted in an ideal world neglecting all kinds of distorting effects from the surroundings such as magnetic fields or detector materials. In general, the event generation is quite fast even for large number ($>10^5$) of events with fairly simple selection criteria.

For the simulation of the CMS detector, a common interface has been written to allow the user to access the most common event generators. In the era before the publication of the physics TDR, the CMKIN software [67] was used as the interface, whereas later the interface has been integrated into the CMS software (CMSSW) [68] framework. The generator interface may be used to apply simple preselection conditions to the generated events to reduce data sample sizes.

In the following, the most common general purpose and dedicated MC event generators are described.

4.1.1 General purpose MC event generators

The most used MC event generator software by the HEP community is PYTHIA [69, 70]. It contains the properties of all the known particles, their production and decay modes, couplings, and interactions. It

uses both theoretical and experimental models to describe a large number of different physics processes such as hard and soft interactions, parton distributions, initial and final state parton showers, multiple interactions, fragmentation, and decay. Some of these descriptions are based on experimental data, but some use theoretical predictions. Besides SM subprocesses, PYTHIA features also subprocesses that can be used to study SUSY as well as other theories beyond the SM such as for example technicolor, leptoquarks or extra dimension models.

Another commonly used MC event generator is ISAJET [71]. It is a Fortran based MC program which simulates pp, $\mathrm{p\bar{p}}$, and $\mathrm{e^+e^-}$ interactions at high energies. It contains perturbative QCD cross-sections, initial and final state QCD radiative corrections in the leading logarithmic approximation, independent fragmentation of quarks and gluons into hadrons and a phenomenological model tuned to minimum bias and hard scattering data for the beam jets. The ISAJET package contains also processes to simulate SUSY or technicolor models.

A third common event generator is HERWIG [72], a Fortran based general purpose MC event generator developed for high energy processes with detailed simulation of QCD parton showers. It simulates hard scattering of hadrons and leptons and the soft hadron-hadron collisions. It uses a parton shower approach for initial and final state QCD radiation and includes color coherence effects and azimuthal correlations both within and between jets. The cluster hadronization of jets is done via non-perturbative gluon splitting. The HERWIG software contains also SUSY models, but lacks other beyond SM processes.

A newer general purpose event generator is MadGraph [73]. It generates automatically the amplitudes for all the relevant subprocesses and produces the mappings for the integration over the phase space which is done with a multi-channel integration method. The process-dependent information is passed to MadEvent and a stand-alone code is produced, which allows the user to calculate cross-sections and to obtain unweighted events. Hence MadGraph may be used as a standalone generator or the generated events can be handed over to any other generator such as PYTHIA or HERWIG for hadronization.

4.1.2 Specialized MC event generators

The general purpose event generators describe usually the processes only up to a non-trivial level of the perturbative expansion. Specialized event generators may be used together with the general purpose ones to obtain next to leading order (NLO) or higher order approximations for the simulation or to provide features, which are not yet included in the general event generators. Some of the specialized event generator packages may also be run as standalone simulations.

For the τ lepton simulation, the most commonly used specialized event generator is TAUOLA [74], which provides simulation of polarized τ lepton decays. Other commonly used packages are ALPGEN [75], which calculates exact matrix elements of hard multiparton processes in hadronic collisions with leading order precision, TopReX [76], which specializes in top quark decays, and HDECAY [77], which calculates the branching ratios and decay widths of SM or MSSM Higgs bosons.

4.2 Detector simulation

The task of the detector simulation is to take the generated events and to impose them to the surrounding environment presented by the detector. To accomplish this, a detailed 3-dimensional model of the studied detector and its magnetic field has to be constructed. The model has to contain information of both the geometry as well as the materials used and their known interaction with different types of particles and radiation. The trajectories of each of the simulated particles, which leave the vacuum of the beam pipe, are then recalculated step by step in the new environment and if an interaction with matter occurs, the production of new particles is also simulated. Finally, the simulated information is digitized as if it would have been measured by the individual subdetectors. The digitization process also includes the realistic modeling of noise in the subdetectors. Additionally, a number of minimum bias events can be superimposed on the hard event to simulate the occurrence of several particle collisions in a single bunch crossing, i.e. pile-up events.

The detector simulation was done at CMS in the era before the publication of the physics TDR with Object oriented Simulation for CMS Analysis and Reconstruction software (OSCAR) [78], which was based on GEANT4 [65] and with CMSIM [79], which was based on GEANT3. In later times, the interface for the GEANT4 has been integrated into the CMSSW framework. The GEANT software simulates the passage of particles through matter. It is used to describe the geometry and materials of the complex detector equipment of modern experiments. It simulates the interaction of the particles passing through the material, tracks the particles, digitizes the response of the particles on the sensitive geometrical volumes and manages the digitized hits in the subdetectors making the output look like the event would have happened in a live detector. The GEANT software provides an extensive set of physics processes, which model the interactions of particles in material volumes. The user may select the desired processes which include electromagnetic, hadronic, and optical processes that are applicable over a large energy range from keV to TeV scales. Besides particle physics, the GEANT package has applications in nuclear physics, heavy ion and radiation calculations, space engineering, and medical physics as well.

The large volume of the tracking of particles and complexity of the detailed simulation requires typically a few minutes of time per event. To save time, GEANT also provides the possibility to run a fast simulation to produce more approximative results before running the full simulation. The results displayed in this thesis are, however, done with full detector simulation.

4.3 Event reconstruction and analysis

The last phase in the simulation chain is to reconstruct physics objects such as jets, tracks, vertices, and further electrons, muons, etc. from the digitized event data and to conduct the final analysis based on the detector information only. In the era before the publication of the physics TDR, the reconstruction and analysis of the events was done with a software called Object-Orientated Reconstruction for CMS

Analysis (ORCA) [80] which was based on the COBRA framework [81]. Later, CMSSW has replaced ORCA and COBRA in the reconstruction and analysis of the events.

Both ORCA and CMSSW are based on the C++ language and take advantage of the modularity of object-orientated programming techniques. Both provide an extensive set of methods to reconstruct the events fully or partially depending on what type of processes are needed for the desired analysis. Furthermore, the reconstructed information can be associated to and compared with the original MC particle information. Since a typical analysis requires several runs over the same reconstructed data samples, the reconstruction of generic physics objects such as jets and tracks as well as the identification of individual particles is done with CMSSW and the information is stored to be used with the physics analysis toolkit (PAT). In such a way, time is saved and the validation of the analysis is simplified.

In the following, the reconstruction of the most essential basic objects, namely jets, tracks, and vertices, is discussed. The description of higher level physics objects, which rely on the properties of the jets, tracks, and vertices, will be described in the later chapters.

4.3.1 Jet reconstruction

Single quarks or gluons, which are in the final state of several processes of interest, cannot be observed directly because of the color quantum number confinement. Instead, these partons form through a process of hadronization a jet of particles whose trajectories are kinematically limited to be within a certain cone. The aim of the jet reconstruction algorithms is to measure with good accuracy the energy of the parton from the energy deposition in the calorimeter cells spanned over a defined portion of space. The reconstruction algorithms are thus required to have good correspondence between the parton, hadronization, and energy deposition levels.

The jets are reconstructed from the energy deposition in the calorimeter towers. In each calorimeter tower, the energy deposition of the HCAL cell and the corresponding ECAL cells are summed. In order to reduce the amount of noise from electronics, a set of thresholds is set on the energy deposition in the calorimeter. The HCAL cell energy deposition must pass an energy threshold of 0.90 GeV and 1.40 GeV in the barrel ($|\eta| < 1.48$) and endcap regions ($1.48 < |\eta| < 3.00$), respectively. A threshold of 1.10 GeV is applied, if the outer barrel cells are active. The sum of the energy deposition of ECAL cells corresponding to the area of a HCAL cell is required to exceed 0.20 GeV and 0.40 GeV in the barrel and endcap regions, respectively. These thresholds correspond to exceeding the root mean square of the expected electronics noise by roughly 2-3 standard deviations of energy deposition.

The jet reconstruction algorithms should be collinear-safe, i.e. the result must be the same, if the energy carried by a single particle is distributed among two collimated particles. Additionally, the jet reconstruction algorithms should be stable against the addition of soft particles (infrared safety). If an algorithm does not satisfy both of these conditions, it will yield ambiguous results and lead to uncertainties in perturbative calculations.

In the following subsections, the three most used jet reconstruction algorithms in CMS, namely the iterative and midpoint cone algorithms and the $\mathrm{k_T}$ algorithm, are presented. The algorithms are followed by the description of the jet resolutions. Finally, an algorithm to calibrate the τ jet energy is presented. Algorithms for jet energy calibration for other than τ jets are described in Ref. [82].

4.3.1.1 Iterative cone algorithm

The iterative cone algorithm [52, 83] starts by sorting calorimeter towers or particles with $\mathrm{E_T} > 1$ GeV. Then, a jet cone of radius $\mathrm{R_{jet}}$ in η,ϕ space is spanned around the object which has the highest $\mathrm{E_T}$. A new direction and energy is calculated for the proto-jet based on the objects inside the cone with the $\mathrm{E_T}$ scheme and the new direction is used as a seed for a new proto-jet. This procedure is repeated until the jet direction in η,ϕ space and the energy of the proto-jet change by less than 1% between the iterations. Once a stable proto-jet is found, it is appended to the list of jets and the objects belonging to it are removed from the object list. The search for jets is repeated, until the $\mathrm{E_T}$ value of the remaining seeds is below a certain threshold.

Since the iterative cone algorithm is fast and since it has predictable execution time, it is used at the high level trigger. The algorithm is neither collinear- nor infrared-safe.

4.3.1.2 Midpoint cone algorithm

The midpoint cone algorithm [84] is based on the iterative cone algorithm. It was designed to allow the splitting and merging of jets. The algorithm searches for jets similarly as the iterative cone algorithm. When a stable proto-jet is found, its constituents are not removed from the object list. Therefore, a single object may be used by several jets and overlapping jet candidates are possible.

In order to address the collinearity and infrared problems, a second iteration is done over the stable proto-jets. For every pair of proto-jets, which are within the diameter of the jet cone, a midpoint is calculated based on the combined momentum. Additional proto-jets are searched for with these midpoints as seeds. Finally, splitting and merging of the proto-jets is carried out starting with the proto-jet with the highest $\mathrm{E_T}$. If the proto-jet with the highest $\mathrm{E_T}$ does not share any objects with another proto-jet, it is appended to the list of jets and it is removed from the proto-jet list. If the proto-jet does contain shared objects, the transverse energy shared with the neighboring proto-jet, which has the highest $\mathrm{E_T}$ of the neighbors, is compared with the total transverse energy of the neighbor proto-jet. If this fraction is greater than a certain value such as 50%, the proto-jets are merged. Otherwise, each shared object is assigned individually to the proto-jet which is closest to the shared object. The splitting and merging procedure is repeated, until no proto-jets are left.

Despite of the improvement on both collinear and infrared safety, the algorithm is infrared-safe only up to next to leading order of perturbative pQCD expansion.

4.3.1.3 Inclusive $\mathrm{k_T}$ algorithm

The inclusive $\mathrm{k_T}$ algorithm [85, 86, 87] is based on clustering. The stable particles or calorimeter cells are used as input objects. For each of these objects i and each of the object pairs (i,j), the distances $\mathrm{d}_i = \mathrm{E}_{\mathrm{T},i}^2$ and $\mathrm{d}_{ij} = \min\{\mathrm{E}_{\mathrm{T},i}^2, \mathrm{E}_{\mathrm{T},j}^2\}\mathrm{R}_{ij}^2\mathrm{D}^2$, where D is a dimensionless parameter and $\mathrm{R}_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2$, are calculated. The algorithm searches for the smallest d_i or d_{ij}. If d_{ij} is the smallest value, then the objects i and j are merged and replaced in the object list as a single object. If d_i is the smallest value, then the object i is declared as a jet and it is removed from the object list. The procedure is repeated until no objects are left in the list. The radius parameter D guarantees the jets to be separated by a distance of $\sqrt{\Delta\eta^2 + \Delta\phi^2} > \mathrm{D}$.

A fast implementation (proportional to $\mathrm{N}\ln\mathrm{N}$) of the $\mathrm{k_T}$ algorithm can be found in Ref. [88]. The $\mathrm{k_T}$ algorithm is collinear- and infrared-safe.

4.3.1.4 Jet resolution

The resolution of the transverse energy of the jets follows the form

$$\frac{\sigma(\frac{\mathrm{E_T^{rec.}}}{\mathrm{E_T^{MC}}})}{\langle\frac{\mathrm{E_T^{rec.}}}{\mathrm{E_T^{MC}}}\rangle} = \frac{\mathrm{A}}{\mathrm{E_T}} \oplus \frac{\mathrm{B}}{\sqrt{\mathrm{E_T}}} \oplus \mathrm{C}. \tag{4.1}$$

The first term describes the fluctuations resulting from electronic noise, pile-up events and underlying event, the second term corresponds to the stochastic response of the calorimeter and the last term is a constant describing the non-uniformities and non-linearities in the detector response. When the equation (4.1) is fitted on jets reconstructed with the iterative cone algorithm ($\mathrm{R_{jet}}$ = 0.5) from simulated QCD multi-jet events with $\hat{\mathrm{p}}_\mathrm{T}$ = 0-4000 GeV/c and $|\eta| < 1.4$, the parameters A = 5.6, B = 1.25 and C = 0.033 are obtained [52].

4.3.1.5 MC-based τ jet energy corrections

The MC-based τ jet energy correction algorithm [89] uses a fit to the MC information to derive calibration curves for τ jets with reconstructed $\mathrm{E_T} > 30$ GeV. The calibration curves are parametrized in 22 bins of jet η as a function of reconstructed jet $\mathrm{E_T}$. The parametrization is carried out with the function

$$\mathrm{E_T(corr)} = \frac{2(\mathrm{E_T} - \mathrm{A})}{\mathrm{B} + \sqrt{\mathrm{B^2} - 4\mathrm{AC} + 4\mathrm{CE_T}}}, \tag{4.2}$$

where A, B and C are calibration constants. The resolution, which can be achieved with this method for τ jets with $\mathrm{E_T} > 30$ GeV and $|\eta| < 2.2$, was found to be $\sigma(\mathrm{E_T})/\mathrm{E_T}$ = 0.057 $\oplus$ 0.932 $/\sqrt{\mathrm{E_T}}$. The statistical uncertainty of the energy correction is $\sim$2%.

4.3.2 Track reconstruction

Charged particles leave a set of hits, i.e. points of impact on the detector layers, as they traverse the tracker or the muon system. The aim of track reconstruction is to reverse engineer the cloud of hits into a set of tracks, i.e. travelling paths of charged particles (track finding). After the hits have been associated to tracks by a pattern recognition algorithm, physical quantities of the tracks, such as momentum and electric charge, are obtained from the track parameters (track fitting). The tasks of identifying hits in the detectors and generating seeds for track finding are not described here, since they are not relevant to the overall topic of this work. A detailed description of these tasks can be found in Ref. [52].

The resolution of the fitted track parameters depend strongly on both the transverse momentum of the track and its pseudo-rapidity. For muons with $\mathrm{p_T} > 100$ GeV/c within $|\eta| < 1.6$, the transverse momentum resolution is found to be $\sim$1-2% and the transverse impact parameter resolution $\sim$10 μm. A detailed description of the performance of the track finding can be found in Ref. [52].

The most commonly used algorithms to perform the track finding and fitting, are the combinatorial Kalman filter (CKF) and Kalman smoother, which are described in Sections 4.3.2.1 and 4.3.2.2, respectively. These algorithms are based on the method of least squares applied on the trajectory parameters and they are excellently suited for reconstructing tracks of charged pions. After the writing of the physics TDR's, more sophisticated track reconstruction algorithms have been developed. One specifically interesting algorithm with respect to this work is the iterative tracking algorithm, which applies track seed generation and track reconstruction in a repetitive manner to increase the performance of track reconstruction. This algorithm is presented in Section 4.3.2.3.

4.3.2.1 Track finding with combinatorial Kalman filter

The combinatorial Kalman filter [90, 91] is a pattern recognition algorithm used for track finding. It uses as input track seeds obtained either from the pixel or strip detectors or both. The seeds provide a rough estimate of trajectory parameters for trajectory candidates. The filter uses these trajectory parameters and extrapolates the track to the next detector layer, where it searches for compatible hits in a region determined by the track parameter accuracy. The effect of magnetic field, multiple interactions, and energy loss resulting from material effects are taken into account in the determination of the search region. For each compatible hit on the new layer, a new trajectory candidate is created. On each detector layer, one trajectory candidate is also created without a hit on the layer to take into account loss of hits in some detector layers.

Each new hit assigned to the trajectory candidate increases the accuracy of the trajectory parameters and thus decreases the search region. The trajectory candidates are grown until they have traversed the whole tracker or until a predetermined number of hits have been found. In order to avoid exponential increase of the number of trajectory candidates as a function of detector layer, some trajectory candidates are discarded after each layer. The decision to keep a trajectory candidate is based on a decision of the

normalized χ^2 of the track fit, the number of compatible hits, and the number of layers in which no hits were found.

4.3.2.2 Track fitting with Kalman filter and smoother

The Kalman filter for track fitting [92, 93, 94] uses the trajectory candidates obtained from the track building as input. The last hit of the trajectory candidates contains the best accuracy of the trajectory parameters, but it might be biased by constraints applied on the seed used for trajectory building. To obtain a good measurement of the track parameters, the trajectory is rebuilt from the hits associated to it starting from the seed, but without constraints on the seed. Such approach increases the accuracy of the trajectory especially in the innermost layers. The track fitting is done with the least squares method.

After the trajectory has been refitted, a smoothing pass is done in the opposite direction. In the smoothing process, the trajectory parameters are updated on each layer taking into account the information of all hits except the one being fitted. The smoothed trajectory can then be used to estimate the track momentum and position at the primary vertex, i.e. the impact point, or at the surface of the ECAL cells. The extrapolation to obtain the estimates is always carried out from the trajectory parameters of the hit closest to the region of interest.

4.3.2.3 Iterative tracking method

The iterative track reconstruction applies the CKF track reconstruction multiple times. In the first track reconstruction, very pure track seeds are used to build tracks with tight vertex cuts and a high minimum number of hits. Therefore, a small fake rate is achieved at moderate tracking efficiency level. The first track reconstruction is followed by three more reconstruction passes. In each of these iteration steps, the hits used in the previous step are cleaned away from the hit pool, new tracks are built with CKF from seeds derived from the remaining hits and the track quality criteria are progressively loosened. This approach results into retaining a low fake rate and recovering a high track reconstruction efficiency. Another benefit of the iterative approach is, that the minimum reconstructable $\mathrm{p_T}$ of the tracks is lowered to $\mathrm{p_T} > 0.3$ GeV/c and the minimum reconstructable number of hits is lowered down to 3.

4.3.3 Vertex reconstruction

Many weakly decaying particles, which are produced in the collision experiments, such as τ's or b hadrons, have a long enough life time to travel a measurable distance before decaying. The task of vertex reconstruction is to use the information of the reconstructed tracks with certain constraints to obtain the position and covariance matrix, i.e. the accuracy of the position, of the production and decaying vertices. The impact point of the two colliding protons is referred to as the primary vertex (PV) and the decay vertices

of observable particles produced in the collision are referred to as secondary vertices (SV). Some decays, such as the decay of b hadrons, may produce even tertiary vertices.

Like track reconstruction, vertex reconstruction consists of the vertex finding and fitting phases. The vertex finding algorithms are designed to group tracks into vertex candidates, after which vertex fitting algorithms are used to obtain the best possible vertex parameters, such as the position and the covariance matrix. The following subsections give an overview of the vertex finding and fitting algorithms used in this work. A more precise description and performance of the vertex reconstructions algorithms is described in Ref. [52].

4.3.3.1 Vertex finding

The main task of vertex finding is the identification of the impact points of the proton-proton collisions in the event. These primary vertex finding algorithms are divided into algorithms usable at the high-level trigger and in offline.

The high-level trigger algorithms for primary vertex finding are based on information only from the pixel detector. Pixel triplets are formed from the hits in the pixel detector with helix approximation [95]. These triplets are then grouped into primary vertex candidates with either the histogramming or divisive method, which are described in Ref. [96]. To save computing time, the primary vertex is assumed to be found on the beam axis. After the vertex finding, the list of primary vertex candidates are sorted in descending order of the sum of the p_T^2 of the triplets associated to them. Only triplets with $\mathrm{p_T} < 10$ GeV/c are taken into account because of the lack of measurement precision. The primary vertex, which belongs to the signal event, is usually the primary vertex candidate which has the highest sum of the p_T^2 of the triplets associated to it.

In offline, primary vertex finding [97] is based on tracks reconstructed in the event. Only tracks which are compatible enough with the beam axis and which are hard enough are taken into account. These tracks are then clustered along the beam axis by grouping tracks, whose impact parameter along the beam axis is separated by less than 1 mm from each other. In each cluster, tracks which are incompatible with the cluster are discarded and the cluster is used for making a fit of a primary vertex candidate. Finally, primary vertex candidates which have bad fits or which are too separated from the beam axis are discarded. Depending on the event topology, primary vertex resolution of 10-40 μm and 15-50 μm are obtained in the transverse plane and in the direction along the beam axis, respectively [52].

The finding of tracks which belong to a secondary vertex is often trivial, since the most robust vertex fitter algorithms accept as input all the tracks within a jet cone. Dedicated algorithms for secondary and tertiary vertex finding in specific physics channels can be found in Ref. [52].

4.3.3.2 Vertex fitting

The most common vertex fitting algorithm is the Kalman vertex fitter (KVF) [98]. It is based on a global least squares minimization. It is the optimal algorithm for linear problems, where the measurements are Gaussian and the fitted parameters depend linearly on the measurements. It is, however, sensitive to outlying measurements, which biases the measurement.

Another algorithm used for vertex fitting is the adaptive vertex fitter (AVF) [99]. It assigns fractional weights on the tracks associated to the vertex based on the standardized distance from the vertex. The weights are varied from one iteration to the next until a converged fit is obtained. The advantage of this algorithm is the capability of applying a soft assignment of the tracks to the vertex.

A third algorithm, which is used for vertex fitting, is the trimmed Kalman fitter (TKF) [100]. It is based on the KVF, but robustified with an algorithm to remove one by one the tracks, which are incompatible with the vertex. The algorithm starts by fitting all input tracks to a vertex. Then, the track least compatible with that vertex is removed, and the vertex is refitted. This procedure is repeated until the compatibility of the remaining tracks to the vertex is above a threshold. The compatibility of each track to the vertex is computed from the standardized distance to the vertex.

4.4 Validation of simulated data

Since the HEP experiments depend on the accuracy of the event generators, detector simulations and the reconstruction of the physics objects, the simulations need to be verified at each simulation level. The event generators can be calibrated with theoretical calculations as well as with experimental data. Typically, distributions of key physics variables produced with different event generators are compared with each other to understand the differences. The detector simulations for both signal and noise are calibrated with the measurement of the individual subdetectors in a test beam. Finally, the reconstruction of physics objects is validated by producing a set of validation data samples and plots of key variables for each new reconstruction software release.

Methods have also been developed to simulate mechanical position uncertainties of the subdetectors. In the following, the effect from the calibration of the sensor positions in the tracker is presented. A more thorough description of the simulation and physics validation procedures used at CMS is presented in Ref. [52].

4.4.1 Simulation of misalignment of tracking sensors

The silicon sensors of the tracker are displaced from their expected positions within mechanical tolerances. These displacements can be compensated with software corrections to each sensor based on data given

by the laser alignment system to align larger structures and on data of reconstructed tracks to align individual sensors [101].

To study the effect of these displacements in the tracker, the misalignment package [102] was used to displace (i.e. to translate and rotate) each sensor individually at the reconstruction level. The displacement was randomly distributed within the tolerances given for each misalignment scenario. For studying the effect of tracker misalignment, the comparison of results in three different misalignment scenarios was proposed. The scenarios correspond to perfect alignment, first data taking, and long-term conditions and are described in the following in more detail.

Scenario 0 represents perfect alignment of the sensors, i.e. no displacements are done. It provides a reference for the best possible results without the effect of sensor displacements.

Scenario 1 is the first data taking scenario, which is expected to correspond to the misalignment conditions during the data taking of the first few hundred pb^{-1} of integrated luminosity. In this scenario, the strip sensors are aligned only with mechanical constraints and the laser alignment system. The pixel detector is expected to have been aligned with a small number of tracks, which has reduced its displacement uncertainties by one order of magnitude from the mechanical tolerances.

Scenario 2 is the long-term scenario, which is expected to correspond to the misalignment conditions after the first few fb^{-1} of integrated luminosity. In this scenario, a complete track-based alignment has been done, which has reduced the strip sensor position uncertainty by a factor of 10 compared to the mechanical constraints. The overall alignment uncertainty for the strip sensors is thus of the order of 20 μm in the long-term scenario.

To account for the sensor displacement in track building, the search cone for the hit on the next layer has to be increased. This was done by adding an alignment position error (APE), specific to each scenario, in quadrature to the hit position error in the denominator of the χ^2 fit of the track and the hit estimates. The introduction of the APE enables thus the recovery of the track reconstruction efficiency, which is lost because of the misalignment, but it also increases the rate of fake tracks. The default APE value given by the misalignment package was used.

5 Selection methods for the $\mathrm{H, A} \to \tau\tau$ and $\mathrm{H}^{\pm} \to \tau^{\pm}\nu_{\tau}$ searches

In this chapter, the methods for reconstructing and for the identification of the higher level physics objects needed in the analysis of the $\mathrm{H, A} \to \tau\tau$ and $\mathrm{H}^{\pm} \to \tau^{\pm}\nu_{\tau}$ channels are presented. The most important event selection method of these is the identification of hadronic τ lepton decays, i.e. τ jets, described in Section 5.1 for τ jets. The other experimental event selection methods exploit the associated b-jet production and the missing $\mathrm{E_T}$ resulting from the neutrino(s). The tagging of the associated b jets and the missing $\mathrm{E_T}$ measurement are described in Sections 5.2 and 5.3. In the $\mathrm{H}^{\pm} \to \tau^{\pm}\nu_{\tau}$ decay channel, the reconstruction of the associated top and W masses described in Section 5.4 is also used.

5.1 Identification of τ jets coming from Higgs boson decays

The efficient and reliable identification of τ leptons is essential for the search for the Higgs boson(s) in different models, as well as the search for many of the so-called new physics phenomena, ranging from SUSY particles to extra dimensions (Table 5.1). Of the τ lepton decays, the hadronic final states are particularly interesting, since they cover $\sim \frac{2}{3}$ of all τ decays and contain only one neutrino.

Channel	Experimental signatures	Application
$\mathrm{gg} \to \mathrm{q\bar{q}H}; \mathrm{H} \to \tau\tau$	jets + τ's + $\mathrm{E_T^{miss}}$	SM Higgs discovery
$\mathrm{gg} \to \mathrm{b\bar{b}H, A}; \mathrm{H, A} \to \tau\tau$	b jets + τ's + $\mathrm{E_T^{miss}}$	Key discovery channel for MSSM H,A
$\mathrm{gg} \to \mathrm{tbH}^{\pm}; \mathrm{H}^{\pm} \to \tau^{\pm}\nu_{\tau}$	b jets + ℓ/jets + τ + $\mathrm{E_T^{miss}}$	Key discovery channel for MSSM $\mathrm{H}^{\pm}$
$\mathrm{q\bar{q}/ggH_1}; \mathrm{H_1} \to \mathrm{A_1A_1} \to 4\tau$	jets + τ's + $\mathrm{E_T^{miss}}$	NMSSM
$\mathrm{H}^{\pm\pm} \to \mu$'s and τ's	μ's + τ's + $\mathrm{E_T^{miss}}$	Doubly charged Higgs (triplet model)
$\tilde{\mathrm{t}}/\mathrm{LQ} \to \tau\mathrm{b}$	b jets + τ's + $\mathrm{E_T^{miss}}$	Stop or third generation leptoquarks
$\phi \to \mathrm{hh} \to \tau\tau\mathrm{b\bar{b}}$	b-jets + τ's + $\mathrm{E_T^{miss}}$	Extra dimensions in the Randall-Sundrum model
$\tilde{\chi}_2^0 \to \tau\tilde{\tau} \to \tau\tau\tilde{\chi}_1^0$	τ's + $\mathrm{E_T^{miss}}$	SUSY cascade

Table 5.1: New physics searches at LHC with τ leptons in the final state.

The basic methods used for the identification of τ jets are very similar, regardless of the decay channel investigated. The τ-jet identification methods presented in this chapter are therefore applicable to several physics channels with a separate optimization of the cut values of the τ identification methods for each physics process. In the following τ-jet identification methods, the focus is on the identification of τ jets from the $\mathrm{H, A} \to \tau\tau$ and $\mathrm{H}^{\pm} \to \tau^{\pm}\nu_{\tau}$ decays, which produce τ's with relatively high $\mathrm{p_T}$. The methods used for the identification of τ jets in the CMS experiment are based on τ properties such as the lifetime, mass, small number of charged decay products, absence of neutral hadrons (neutrons and

$\mathrm{K^0}$'s), and collimated and isolated decay signature. The usage of these methods in different combinations depends on the physics channel under consideration. In the case of the identification of the τ's from the $\mathrm{H^{\pm}} \to \tau^{\pm}\nu_{\tau}$ decay against the τ's from the $\mathrm{W^{\pm}} \to \tau^{\pm}\nu_{\tau}$ decay, τ helicity correlations can also be exploited.

In the following, the signal and background samples used to evaluate the τ-jet identification methods are described in Section 5.1.1. The kinematical properties of τ jets are presented in Section 5.1.2. The Level-1 and high level single and double τ triggers are described in Sections 5.1.3 and 5.1.4, respectively. Sections 5.1.5-5.1.12 cover the individual τ-jet identification methods. In Section 5.1.13, the effect of tracker misalignment on tracker based τ-jet identification methods and the plans for the calibration of the τ-jet identification with data are discussed. Finally, Section 5.1.14 summarizes the recent developments in the τ-jet identification methods after the publication of the CMS physics TDRs. The final optimization of the τ-jet identification methods is done in the discussion of the $\mathrm{gg} \to \mathrm{b\bar{b}H(A)}$, $\mathrm{H, A} \to \tau\tau$ and $\mathrm{gg} \to \mathrm{tbH^{\pm}}$, $\mathrm{H^{\pm}} \to \tau^{\pm}\nu_{\tau}$ decay channels in chapters 6 and 7, respectively.

All distributions and efficiencies shown in the following subsections, except for Sections 5.1.4 (high level trigger), 5.1.11 (neutral hadron rejection), and 5.1.12 (helicity correlations), are relative to events passing a MC preselection. The preselections consists of matching the direction of reconstructed jets with the direction of the MC τ or jet direction within a cone of 0.2 in η,ϕ space and of requiring the reconstructed jets to be within $|\eta| < 2.2$. Furthermore, the two hadronic jets from QCD multi-jet events were required to be separated in η,ϕ space by $\Delta\mathrm{R} > 1.5$

5.1.1 Signal and background samples

The τ events used to evaluate the τ-jet identification methods were generated with the PYTHIA [69] package. In these events, the τ leptons were produced back-to-back with the same momentum and they were forced to decay hadronically. The transverse energy of the generated τ leptons was generated with a flat range of 20 to 420 GeV and a uniform distribution in ϕ. The pseudo-rapidity of the τ leptons was limited to $|\eta| < 2.2$. No pile-up or underlying event was simulated for the signal unless otherwise stated. For the evaluation of the Level-1 and high level trigger, events were generated with the PYTHIA [69] process $\mathrm{gg} \to \mathrm{b\bar{b}H(A)}$, $\mathrm{H, A} \to \tau\tau$. The $\mathrm{m_H}$ points of 200, 500, and 800 GeV/c^2 were simulated with both τ leptons forced to decay hadronically. Pile-up events corresponding to the LHC low-luminosity scenario of $\mathcal{L} = 2{\times}10^{33}$ cm^{-2}s^{-1} were included in the simulation of the $\mathrm{H, A} \to \tau\tau$ events.

The most dangerous background for the τ jets are the hadronic jets from QCD multi-jet events, because the cross-section of the QCD multi-jet events exceeds the cross-section of electroweak backgrounds such as $\mathrm{t\bar{t}}$ and W+jets by a considerable factor. The background rejection with the τ-jet identification methods was evaluated for a sample of QCD multi-jet events generated with PYTHIA [69] (MSEL=1) with transverse energy of the jets between 30 and 170 GeV in bins of 30-50, 50-80, 80-120, and 120-170 GeV. Higher jet energies were not considered for the QCD multi-jet events, since their charged track multiplicity is on

the average much higher than that of the τ jets and since their production cross-section decreases rapidly with jet $\mathrm{E_T}$. Pile-up events corresponding to the luminosity of $\mathcal{L} = 2{\times}10^{33}$ cm^{-2}s^{-1} were included in the background simulation. The two MC jets with the highest $\mathrm{E_T}$ in the QCD multi-jet sample were considered as τ-jet candidates. These two jets were required to be separated in η,ϕ space by the distance of $\Delta\mathrm{R} > 1.5$, and to be reconstructed inside $|\eta| < 2.1$, after which they were propagated through the τ-jet identification criteria.

The τ-jet identification performance was evaluated as a function of the true transverse energy $\mathrm{E_T^{MC}}$ and the pseudo-rapidity of the jet. The true τ-jet transverse energy was defined as the MC energy of the τ lepton without the neutrino energy in the decay $\tau \rightarrow \mathrm{hadrons}+\nu$. The true energy of hadronic jets from QCD multi-jet events was defined as the energy of stable MC particles, excluding neutrinos and muons, in a cone of 0.5 around the reconstructed jet axis. The efficiency for passing the MC preselection criteria was found to be of the order of 12% for the QCD multi-jet events.

The detector simulation was carried out with GEANT3 [103] in a magnetic field strength of 4 T corresponding to the nominal design value of the CMS solenoid. The reconstruction and analysis of the events was done with the ORCA software version 8.7.4. Jets were reconstructed with the iterative cone algorithm described in Section 4.3.1.1 in a cone of 0.4. The energy deposition of calorimeter towers was required to fulfill $\mathrm{E_T} > 0.5$ GeV and $\mathrm{E} > 0.8$ GeV. Tracks were reconstructed with the combinatorial Kalman filter and Kalman smoother described in Sections 4.3.2.1 and 4.3.2.2. For the signal events, the simulated primary vertex z coordinate ($\mathrm{PV_z}$), i.e. the position of the interaction point along the beam axis, was taken as the reconstructed $\mathrm{PV_z}$ with a Gaussian smearing of 60 μm. The x and y coordinates of the primary vertex were taken to be zero in the plane transverse to the beam axis with a Gaussian smearing of 10 microns. For the background, the primary vertex position was reconstructed with the method described in 4.3.3.1 in three-dimensional space with a resolution of better than 35 μm for the z coordinate.

The τ-jet identification for the $\mathrm{H^{\pm}} \rightarrow \tau^{\pm}\nu_{\tau}$ decay was studied with genuine $\mathrm{gg} \rightarrow \mathrm{tbH^{\pm}}$, $\mathrm{H^{\pm}} \rightarrow \tau^{\pm}\nu_{\tau}$ events without pile-up events. The generation of these event samples is described in Section 7.7.1.1. These event samples were used for the neutral hadron rejection method (Section 5.1.11) and the helicity correlation discussion (Section 5.1.12). The events were reconstructed with the CMSSW [68] software version 1.6.12. The iterative cone algorithm (Section 4.3.1.1) with a jet cone of 0.5 and the iterative tracking method (Section 4.3.2.3) were used for jet and track reconstruction in the events, respectively.

5.1.2 Kinematical τ-jet properties relevant for τ-jet identification

The branching ratios of the τ leptons are described in detail in Section 2.3.1. The τ lepton decays hadronically 64% of the time with a relatively small number of charged and neutral hadrons in the final state. In 75% of hadronic τ decays, the τ jet consists of only one charged hadron and zero or more π^0's (one-prong decays). Furthermore, in 23% of hadronic τ decays, the τ jet consists of three charged

hadrons and zero or more π^0's (three-prong decays). Decay modes with five or more charged hadrons as well as K^0's exist, but their branching ratio can be considered to be negligible. In both one- and three-prong final states the reconstructed charged track with the highest $\mathrm{p_T}$ at origin is labeled as the leading track. When the $\mathrm{p_T}$ of the τ jet is large compared to the τ mass (m_{τ} = 1.78 GeV/c^2), these hadrons have relatively small momentum in the plane transverse to the τ jet axis because of the boost effect. Hence, a hadronic τ decay produces a narrow jet-like cluster in the calorimeter, which is why the hadronic τ lepton decays have been coined as τ jets.

The ratio of transverse energy of the reconstructed jet compared to transverse energy of the MC generator-level jet, $\mathrm{r} = \mathrm{E_T^{reco}}/\mathrm{E_T^{MC}}$, and the transverse energy resolution of the τ jet, are shown respectively in the left and right plots of Fig. 5.1 as a function of the reconstructed jet cone size in η,ϕ space for three bins of true τ-jet energy. The values of r in the left plot of Fig. 5.1 were normalized to the value obtained with a jet reconstruction cone size of 0.6. Based on Fig. 5.1, a jet reconstruction cone size of at least 0.4 was found to contain more than 98% of the τ-jet energy. Furthermore, a cone size smaller than 0.4 was observed to deteriorate the τ-jet energy resolution. Therefore, a jet reconstruction cone size of 0.4 or even 0.5 was chosen for the analyses in the physics TDRs. The jet reconstruction cone size of 0.5 was used in the neutral hadron rejection method and helicity correlation discussion as well as the reoptimization of the τ-jet identification for the $\mathrm{H}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ channel in Section 7.7.1. A larger jet reconstruction cone size than 0.5 can lead to a contamination from other jets in multi-jet events and from particles from pile-up events.

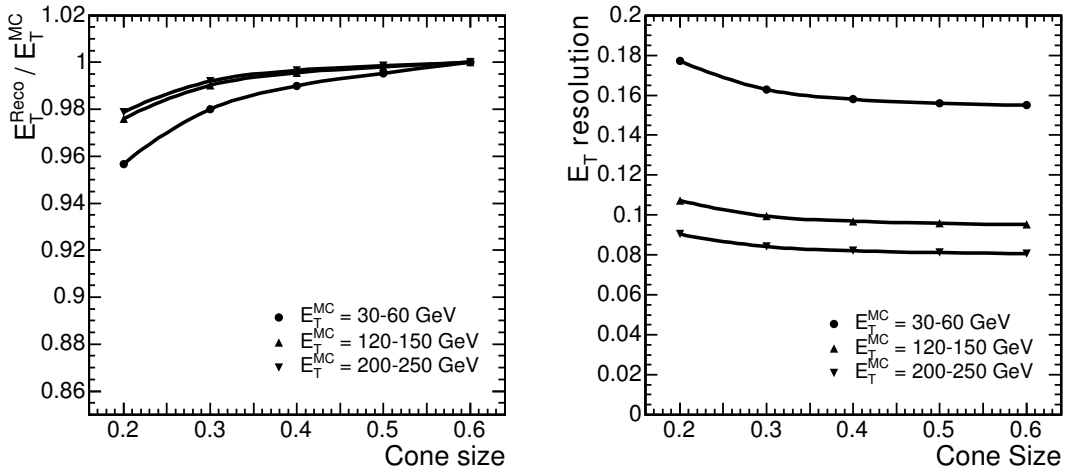

Figure 5.1: Distribution of the ratio $\mathrm{r} = \mathrm{E_T^{reco}}/\mathrm{E_T^{MC}}$ (left) and the transverse energy resolution (right) as a function of the jet reconstruction cone size for the three bins of τ jet $\mathrm{E_T^{MC}}$. In the left plot, the values of r were normalized to the r for a jet reconstruction cone size of 0.6.

The difference in the azimuthal angle (ϕ) and pseudo-rapidity (η) directions between the true and reconstructed τ-jet is shown in Fig. 5.2 for three intervals of the true τ-jet energy. The charge of the τ lepton was positive in these event samples. The 4 Tesla magnetic field can be seen to cause a systematic shift of 0.02 radians in the reconstructed τ-jet direction in ϕ for τ jets with $\mathrm{E_T^{MC}}$ between 40 and 60 GeV. The shift was found to diminish for the τ jets with larger $\mathrm{E_T^{MC}}$. The difference in η direction was found to be

independent of the true τ jet transverse energy for τ jet with $\mathrm{E_T^{MC}}$ between 40 and 250 GeV and slightly worse than in ϕ direction.

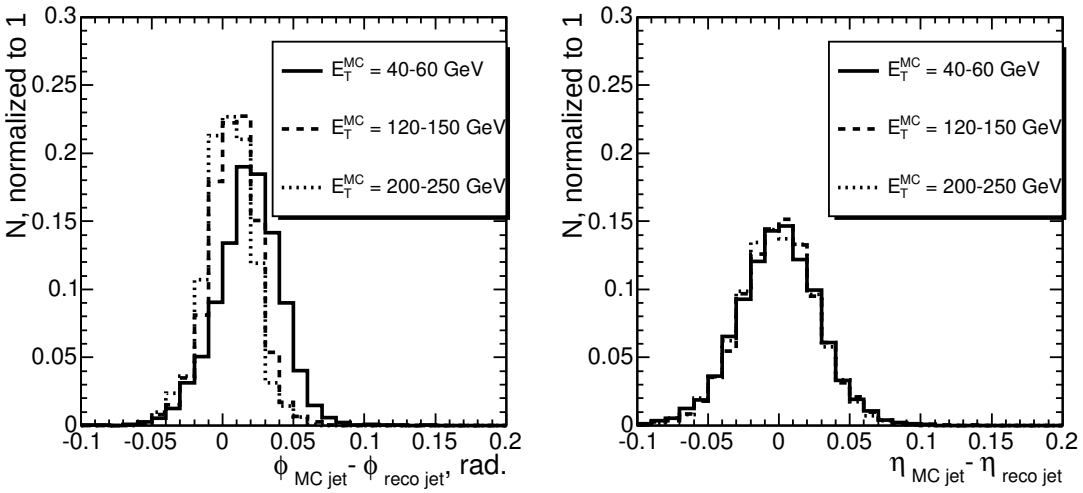

Figure 5.2: Distribution of the difference in ϕ (left plot) and in η (right plot) between the true τ-jet direction and the jet direction reconstructed with the calorimeter for the three different intervals of the true τ-jet energy. The τ lepton has a positive charge in these event samples.

A matching between the calorimeter jet axis and the charged particles from the hadronic τ decays measured with the tracker was required. The distance $\Delta\mathrm{R} = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ in η,ϕ space between the leading track direction and the direction of the reconstructed τ jet are shown in Fig. 5.3 for three bins of the true τ-jet transverse energy. A cut of 3 GeV/c was applied on the $\mathrm{p_T}$ of the leading track in the figure. Both, the one- and the three-prong τ decays were included. It can be seen in Fig. 5.3 that the value of $\Delta\mathrm{R}$ does not exceed 0.1 for the range of τ-jet $\mathrm{E_T^{MC}}$ considered.

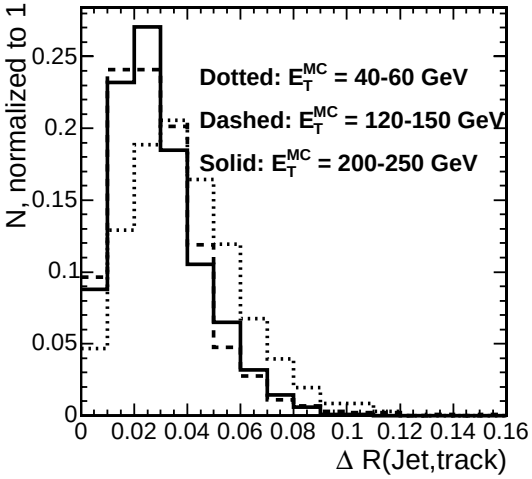

Figure 5.3: Distribution of the distance, $\Delta\mathrm{R}$, in η,ϕ space between the reconstructed leading track and the direction of the reconstructed τ jet for three bins of the true τ-jet transverse energy including both one- and three-prong τ decays. The cut of 3 GeV/c was applied on the $\mathrm{p_T}$ of the leading track.

In three-prong τ decays, the three charged hadrons are produced within a narrow cone because of the

boost effect. The maximal distance $\Delta\mathrm{R}$ in the η,ϕ space between the leading track and the other two tracks of the three-prong τ decays is shown in Fig. 5.4 for three bins of the true τ-jet transverse energy. It can be seen in Fig. 5.4 that in more than 95% of the three-prong τ decays with τ-jet $\mathrm{E_T^{MC}}$ above 60 GeV, the tracks are contained within a cone of 0.07 around the leading track in η,ϕ space. The cone necessary to contain the tracks shrinks with τ-jet $\mathrm{E_T^{MC}}$ because of the boost effect. Therefore, a cone of 0.04 around the leading track was found to contain more than 95% of the tracks from three-prong τ decays with τ-jet $\mathrm{E_T^{MC}}$ above 100 GeV.

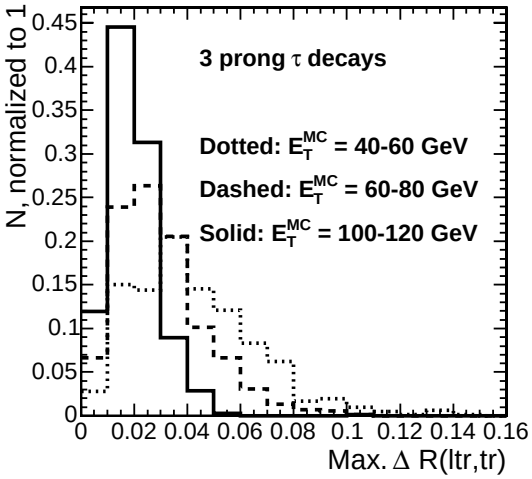

Figure 5.4: Maximal distance $\Delta\mathrm{R}$ in η,ϕ space between the leading track and the other two charged particles from in three-prong τ decays for three bins of the true τ-jet transverse energy.

5.1.3 First level trigger

The first level (Level-1) trigger algorithm is very similar for both hadronic jets and τ jets [62]. The algorithm starts by forming trigger regions from the η,ϕ grid of calorimeter trigger towers (see Section 3.2.2) depicted by the squares in Fig. 5.5. In the central and endcap regions of the calorimetry, trigger regions are formed of blocks of 4×4 trigger towers marked by thick black lines in Fig. 5.5. In the forward calorimeters, where the η and ϕ granularity is coarser, a single trigger tower forms a trigger region. In such a way, the trigger regions are used to cover the whole η,ϕ grid.

The algorithm scans the η,ϕ space with a sliding window consisting of 3×3 trigger regions and analyses the transverse energy deposition in them. A Level-1 jet is found, if the transverse energy deposition is higher in the central trigger region of the sliding window than in any of the eight surrounding trigger regions, and if the transverse energy deposited in the central region of the sliding window exceeds a certain $\mathrm{E_T}$ threshold. In order to separate hadronic jets from τ jets, the narrowness of the energy deposition pattern of hadronic τ decays is exploited. This pattern search is carried out by examining the energy deposition in the trigger towers separately for each trigger region of the sliding window. The trigger towers are labeled active, if their ECAL or HCAL energy deposition exceeds a certain programmable

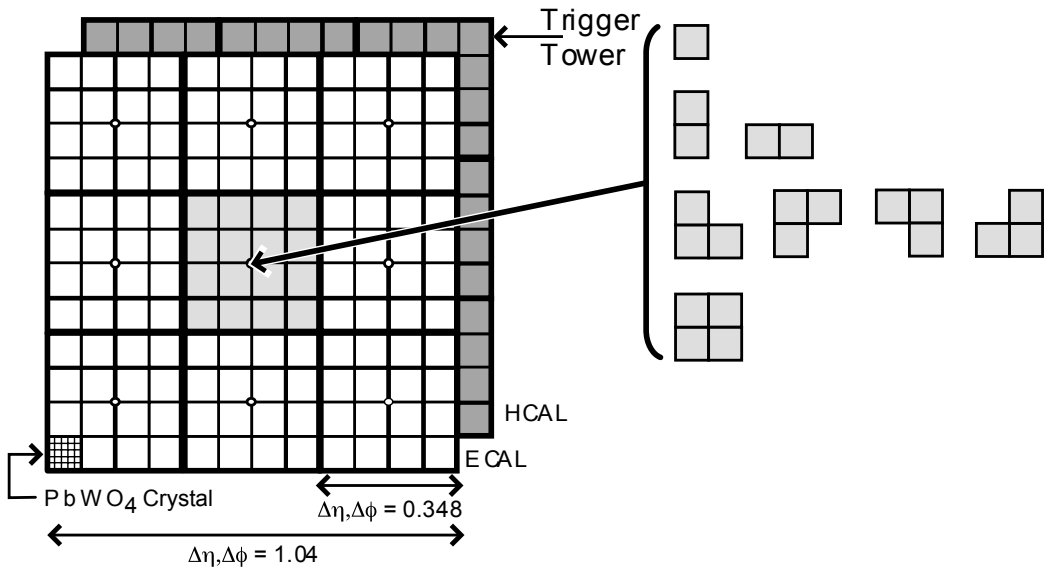

Figure 5.5: Visualization of hadronic jet and τ-jet Level-1 trigger algorithms from Ref. [62].

threshold. If the active towers in a trigger region do not match any of the allowed patterns shown on the right in Fig. 5.5, they are assigned a τ veto bit. If none of the nine τ veto bits of the trigger regions in the sliding window are set, then the Level-1 jet is labeled as a Level-1 τ jet.

By scanning the calorimeter η,ϕ space with the sliding window, the Level-1 jet algorithm produces a list of Level-1 jets and another list of Level-1 τ jets, which are used as trigger primitives. Both lists are sorted in descending order as a function of jet $\mathrm{E_T}$. In order to trigger events with hadronic jets or τ jets, a certain number of Level-1 hadronic and/or τ jets are required to have an $\mathrm{E_T}$ which exceeds a given threshold. For the $\mathrm{H}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ and the $\mathrm{H, A} \rightarrow \tau\tau$ decay channels, the single τ and double τ triggers, which require one or two Level-1 τ jets exceeding a certain jet $\mathrm{E_T}$ threshold in each event, respectively, are the most interesting triggers.

The background, which has by far the highest cross-section and which consequently needs to be suppressed most with the trigger, is the QCD multi-jet background. The rate of QCD multi-jet events passing the Level-1 single, double and single or double τ triggers is shown in Table 5.2 for the luminosity of 2×10^{33} $\mathrm{s^{-1}cm^{-2}}$. The event rates were obtained with the jet $\mathrm{E_T}$ thresholds of 93 and 66 GeV for the Level-1 τ jets for the single and double τ triggers, respectively. It can be observed, that the dominant contribution to the event rate comes from the three $\hat{\mathrm{p}}_{\mathrm{T}}$ bins in the interval between 50 and 170 GeV/c.

5.1.4 High level trigger

The high level trigger (HLT) [62] for τ jets is based on a $\mathrm{p_T}$ cut on the leading track of a jet and on electromagnetic and charged track isolation, which are described in detail in Sections 5.1.5 and 5.1.6, respectively. These methods were found to reduce the QCD background to a tolerable level for the $\mathrm{H, A} \rightarrow \tau\tau$

$\hat{\mathrm{p}}_{\mathrm{T}}$, GeV/$c$	Cross-section, fb	Trigger rate, kHz single τ	double τ	single/double τ
30-50	1.56×10^{11}	0.04	0.08	0.12
50-80	2.09×10^{10}	0.59	0.70	1.19
80-120	2.94×10^{9}	1.32	0.75	1.65
120-170	5.00×10^{8}	0.46	0.16	0.48
170-230	1.01×10^{8}	0.10	0.03	0.10
230-300	2.39×10^{8}	0.02	0.007	0.021
Total rate		2.53	1.73	3.56

Table 5.2: Rate for the QCD multi-jet background in kHz at a luminosity of 2×10^{33} $\mathrm{s^{-1}cm^{-2}}$ for the Level-1 single, double, and single or double τ triggers with single (double) trigger threshold of 93 (66) GeV.

channel. For the $\mathrm{H}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ channel, an additional cut has to be applied on the uncorrected missing transverse energy measurement described in Section 5.3.

At the high level trigger, the iterative cone algorithm described in Section 4.3.1.1 is used to search for jets within a cone of 0.8 around the one or two Level-1 τ-jet candidate directions and to reconstruct the jets in a jet cone of 0.6. If two τ jets are asked for and if only one Level-1 τ jet is found, the generic Level-1 jet with the highest transverse energy was taken as the second HLT τ jet candidate. For the signal events, the two jets selected in this way had a purity of 97% and 82% for the first and second HLT τ-jet candidate, respectively. The purity was found to be independent of the $\mathrm{m_A}$ mass.

After the reconstructions of the jets, a light τ-jet identification is applied on the HLT τ-jet candidates. The following approaches were investigated:

- electromagnetic isolation, followed by charged track isolation with the tracks reconstructed only with the pixel detector (Calo+Pxl) and
- charged track isolation with regional track reconstruction based on hits from both the pixel and the silicon tracker layers (Trk-tau).

The first approach is fast and gives a good performance as far as the electromagnetic and charged track isolation algorithms are concerned. It is therefore the preferred approach for decays with two τ jets in the final state, where the electromagnetic and charged track isolation suppress sufficiently the backgrounds such as in the $\mathrm{H, A} \rightarrow \tau\tau$ decay channel. The second approach is slower but gives a more accurate estimation of the charged track momenta. It is therefore useful in decay channels such as $\mathrm{H}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$, where the high level trigger selection contains a large $\mathrm{E_T^{miss}}$ in the event, and where in addition to the charged track isolation a tight cut on the momentum of the leading track $\mathrm{p_T}$ is required [62]. More details on the logic of the trigger system can be found in Refs. [62, 104, 105, 106].

The HLT efficiencies were evaluated for the $\mathrm{gg} \rightarrow \mathrm{b\bar{b}H(A)}$, $\mathrm{H, A} \rightarrow \tau\tau$ process with two τ jets in the final state. The possible overlapping of an associated b jet with a τ jet can cause the τ jet to appear

unisolated, which can reduce the trigger performance in the signal events. The QCD multi-jet events with $\hat{\mathrm{p}}_{\mathrm{T}}$ bins of 50-170 GeV/c were used as background, since they were found to give the most contribution to the event rate at the Level-1 trigger. The efficiencies presented for the HLT approaches are given with respect to events which pass the Level-1 single or double τ trigger.

5.1.4.1 The high level τ trigger based on em. and pixel track isolation

The electromagnetic plus the pixel-track isolation at the high level trigger is referred to as the Calo+Pxl τ trigger. In this approach, the electromagnetic isolation described in Section 5.1.5 is applied to the τ jet with the highest transverse energy. The efficiency of this cut with respect to the Level-1 single or double τ trigger is shown in Fig. 5.6 as a function of $\mathrm{P}_{\mathrm{isol}}^{\mathrm{cut}}$. It can be seen from the figure, that the selected cut point of $\mathrm{P}_{\mathrm{isol}}^{\mathrm{cut}} < 5$ GeV yields a rejection factor of $\sim$3 against the QCD multi-jet background.

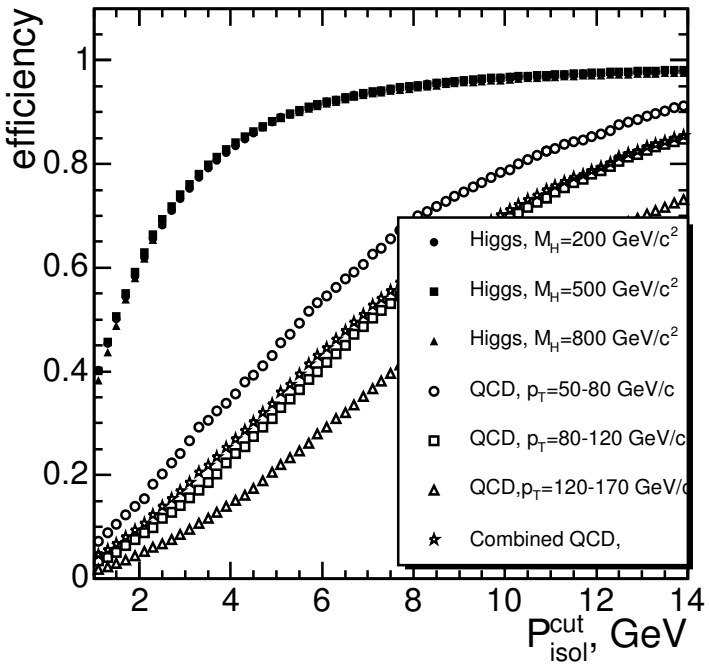

Figure 5.6: Efficiency of the electromagnetic isolation at the high level trigger for the signal and for the QCD multi-jet background as a function of the cut $\mathrm{P}_{\mathrm{isol}}^{\mathrm{cut}}$ on the electromagnetic isolation parameter.

The remaining background rate is suppressed with charged track isolation based on information from the Pixel detector alone. The charged track isolation requirement is applied to both jets. The algorithm relies on finding three pixel hits per track to reconstruct the pixel-tracks [107]. The pixel-tracks are built from seeds of pixel hit pairs from the first two layers of the pixel detector (barrel+barrel or barrel+endcap). These hits are matched in r-ϕ and z-r planes to create the pixel-track seeds. Valid pixel pairs are then matched with a third pixel hit forming pixel-tracks. The momentum of the pixel-tracks is then reconstructed from the three pixel hits without the primary vertex constraint. The number of fake pixel-tracks in

the isolation cone was found to be very low (3-4%). The cuts were optimized for a minimum pixel-track $\mathrm{p_T}$ of 1 GeV/c.

The z coordinate of the primary vertex is then reconstructed from the pixel-tracks. A list of primary vertices is formed from those z values in which two or more tracks cross the z axis. Only primary vertices with at least three valid pixel-tracks are kept and the position of each vertex is estimated as the mean value of the z impact parameters of all pixel-tracks assigned to it [107]. The primary vertex position is then defined to be the primary vertex candidate which is associated to the leading track from the jet with the highest energy. This selection was found to yield in 99% of the signal events the same reconstructed primary vertex z position as the simulated vertex. Pixel-tracks associated to this primary vertex are then considered for the charged track isolation algorithm and tracks associated to other primary vertex candidates are ignored.

The charged track isolation applied to the pixel-tracks is the same as the algorithm described for full tracks in Section 5.1.6. The leading pixel-track was required to be found within a matching cone of $\mathrm{R_m} = 0.1$ around the HLT τ-jet axis and the transverse momentum of the leading pixel-track was required to exceed 3 GeV/c. The signal cone size $\mathrm{R_s}$ was set to 0.07 and the isolation cone size $\mathrm{R_i}$ was varied as a free parameter to adjust the trigger rate.

The performance of the Calo+Pxl trigger is shown in Fig. 5.7 for the HLT τ jet with the highest energy (left plot) and for both HLT τ jets (right plot). The nine evaluation points in the figure correspond to steps of 0.05 of the isolation cone $\mathrm{R_i}$ between 0.2 and 0.6. The single τ trigger was found yield a rejection factor of $\sim$30 can be achieved against the QCD multi-jet background with the isolation cone size of $\mathrm{R_i} = 0.40$ while retaining a signal efficiency of $\sim$54-58%. The double τ trigger was found to suppress the QCD multi-jet background by a factor of about 10^3 with the isolation cone size $\mathrm{R_i}$ between 0.45 and 0.50 while retaining a signal efficiency of 29-32%.

5.1.4.2 The high level τ trigger based on charged track isolation

The high level trigger algorithm based on charged track isolation is referred to as the Trk-Tau trigger. Because of the time limitation at the high level trigger, it is not possible to perform a full track reconstruction of the whole event after the Level-1 single or double τ trigger. It is, however, possible to read and reconstruct a selected part of the data provided by the tracker. Therefore, regional tracking is used in the Trk-Tau approach to reconstruct only the tracks in a cone around the HLT τ jet directions. To save time, the track finding is stopped for each track once they contain seven hits.

The primary vertex is obtained in this algorithm from information of the pixel detector alone to ensure fast reconstruction. Only tracks, which have a z impact parameter close enough to the z position of the primary vertex, are considered for the charged track isolation. The leading track was required to be found within a matching cone of $\mathrm{R_m} = 0.1$ around the HLT τ jet. The transverse momentum of the leading track was required to exceed 6 GeV/c. The signal cone size was set to $\mathrm{R_m} = 0.07$. The size of the isolation

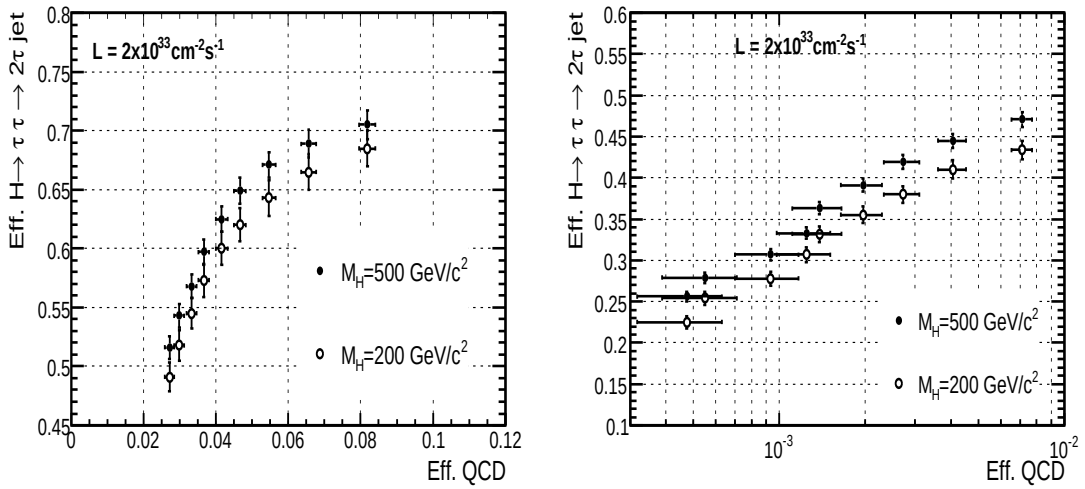

Figure 5.7: Efficiency of the Calo+Pxl trigger trigger applied to the first HLT τ jet (left plot) and to both HLT τ jets (right plot) for signal events versus QCD multi-jet events for the Higgs boson masses of $\mathrm{m_H}$ = 200 and 500 GeV/c^2. The dots correspond to isolation cone size $\mathrm{R_i}$ values between 0.2 to 0.6 varied in steps of 0.05.

cone $\mathrm{R_i}$ was treated as a free parameter and varied between 0.2 and 0.45 in steps of 0.05.

The Trk-Tau trigger performance is shown in Fig. 5.8 for the HLT τ jet with the highest energy (left plot) and both HLT τ jets (right plot). It can be seen from the figure, that a rejection factor of $\sim$20 ($\sim 10^3$) can be achieved against the QCD multi-jet background with the single (double) τ trigger with an isolation cone size $\mathrm{R_i}$ of about 0.40, while retaining a signal efficiency of $\sim$65-75% ($\sim$35-40%).

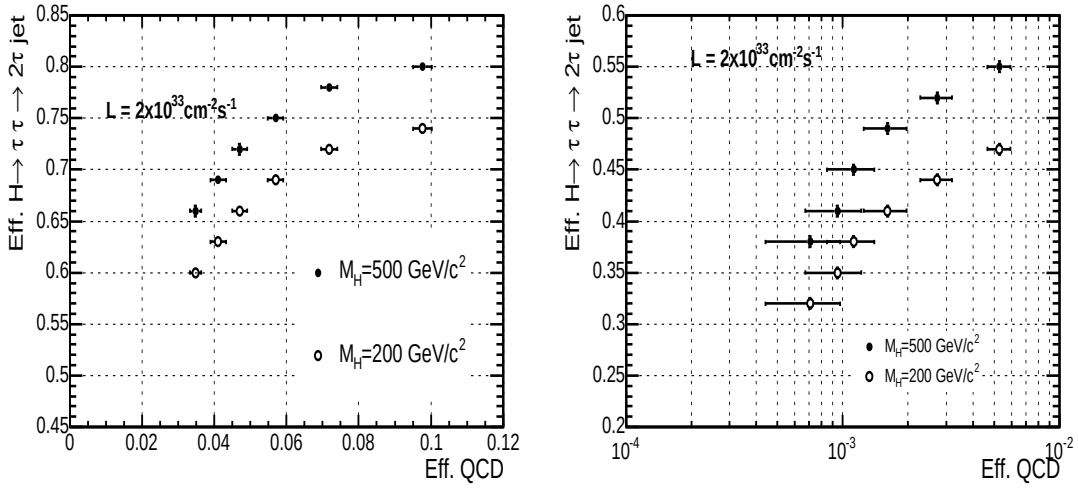

Figure 5.8: Performance of the Trk-Tau trigger applied to the HLT τ jet with the highest energy (left plot) and to the two HLT τ jets (right plot) against the QCD multi-jet background for the Higgs boson masses of $\mathrm{m_H}$ = 200 and 500 GeV/c^2. The points correspond to isolation cone size $\mathrm{R_i}$ values between 0.2 and 0.45 varied in steps of 0.05.

The main difference between the performances of the Trk-Tau and Calo+pxl algorithm comes from the better momentum resolution of the tracks in the Trk-Tau approach compared to the momentum resolution

of the pixel-track in the Calo+Pxl approach. The better momentum resolution enables a stronger cut on the $\mathrm{p_T}$ of the leading track.

5.1.5 Electromagnetic isolation

Hadronic τ decays produce a localized energy deposit in the electromagnetic calorimeter. The source of this electromagnetic energy deposit are the π^0's, which are produced in the τ decay, and which decay subsequently into a photon pair. Additionally, the showering process of charged hadron(s) in the τ jets can start already in the electromagnetic calorimeter (ECAL) contributing hence to the electromagnetic energy deposit. Because of the boost effect, the electromagnetic energy of a τ jet is typically contained within a narrow signal cone around the leading track direction.

Several variables were tried to quantify this feature and to use it for τ-jet identification and rejection of hadronic jets from QCD multi-jet events [104]. The electromagnetic isolation parameter $\mathrm{P_{isol}}$ defined as

$$\mathrm{P_{isol}} = \sum_{\mathrm{R_2}} \mathrm{E_T^{em.}} - \sum_{\mathrm{R_1}} \mathrm{E_T^{em.}} \tag{5.1}$$

was found to provide the best performance. The transverse energy deposits in the electromagnetic calorimeter is summed in a signal cone of radius $\mathrm{R_1}$ and isolation cone of radius $\mathrm{R_2}$ in η,ϕ space around the leading track direction. The energy deposit in the signal cone is then substracted from the energy deposit in the isolation cone to obtain the quantity $\mathrm{P_{isol}}$. Jets with $\mathrm{P_{isol}} < \mathrm{P_{isol}^{cut}}$ are considered as τ-like. More information about the choice of this variable and the parameters can be found in Refs. [108, 109].

The efficiency of the electromagnetic isolation is shown in Fig. 5.9 for τ jets as a function of $\mathrm{E_T^{MC}}$ (left plot) and $|\eta^{\mathrm{MC}}|$ (right plot) for $\mathrm{R_2}$ = 0.13, $\mathrm{R_2}$ = 0.40 and $\mathrm{P_{isol}^{cut}} = 5$ GeV. The efficiency is shown separately for the four final states of hadronic τ lepton decays. Only a small ($\sim$5%) variation was observed in the signal efficiencies as a function of $\mathrm{E_T^{MC}}$ over a large region of simulated transverse energies from 30 to 300 GeV. The variation in pseudorapidity for τ decays with π^0's in the final state follows the variation of the amount of the tracker material in front of the ECAL. This correlation results from electrons and positrons from photon conversions in the tracker material contaminating the electromagnetic isolation region.

The performance of the electromagnetic isolation is shown in Fig. 5.10 for τ jets and hadronic jets from QCD multi-jet events in different $\mathrm{E_T^{MC}}$ bins as a function of the $\mathrm{P_{isol}^{cut}}$ value. It can be seen in the figure, that the electromagnetic isolation can provide a rejection factor of $\sim$5 against hadronic jets from QCD multi-jet events with jet $\mathrm{E_T^{MC}} > 80$ GeV with a signal efficiency of better than 80%. The efficiency for the hadronic jets from QCD multi-jet events was found to decrease as the $\mathrm{E_T^{MC}}$ of the jet increases. This behavior is caused by low energy charged particles ($\mathrm{p_T} < 2$ GeV/c), which bend out of the cone of 0.4 and do thus not contribute to the energy sum in the $\mathrm{P_{isol}^{cut}}$ formula [110].

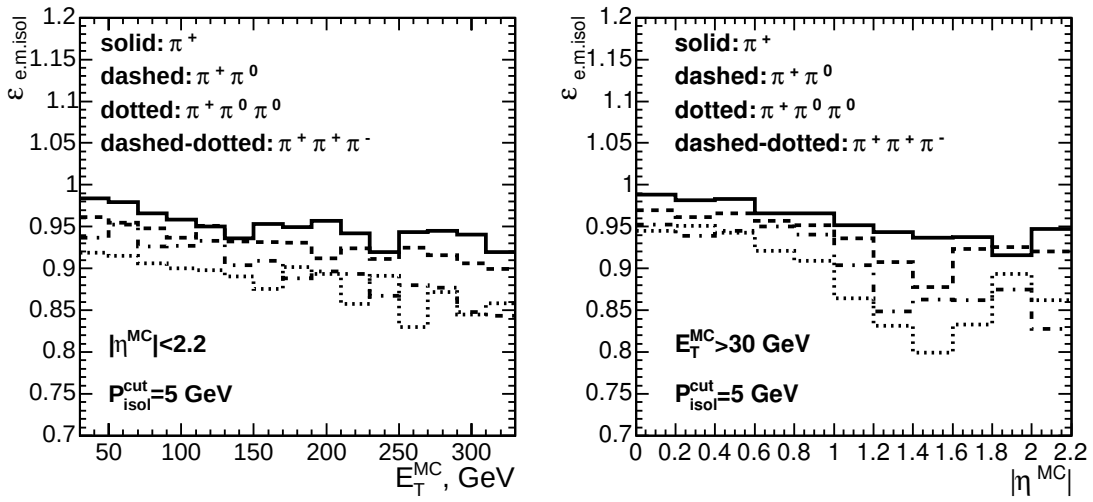

Figure 5.9: Efficiency of the electromagnetic isolation for τ jets as a function of $\mathrm{E_T^{MC}}$ (left plot) and $|\eta^{\mathrm{MC}}|$ (right plot) for $\mathrm{P_{isol}^{cut}} = 5$ GeV. The efficiency is shown separately for several final states of hadronic decays of τ lepton.

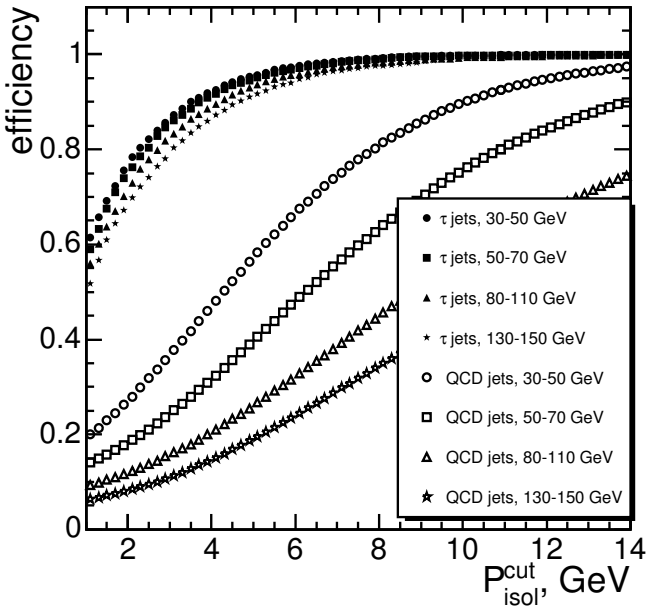

Figure 5.10: Efficiency of the electromagnetic isolation for τ jets and hadronic jets from QCD multi-jet events for the different bins of the true simulated transverse energy as a function of the $\mathrm{P_{isol}^{cut}}$.

5.1.6 Charged track isolation

The principle of τ-jet identification with the charged track isolation is visualized in Fig. 5.11. The direction of the τ jet is defined by the axis of the reconstructed jet. The leading track is required to be found within a matching cone of radius $\mathrm{R_m}$ around the calorimeter jet direction. Around the leading track direction, signal and isolation cones of radii $\mathrm{R_s}$ and $\mathrm{R_i}$, respectively, are spanned. The tracks inside the signal cone and the isolation annulus between the signal and isolation cones are counted. To apply the charged track

isolation condition, no reconstructed tracks are allowed inside the isolation annulus.

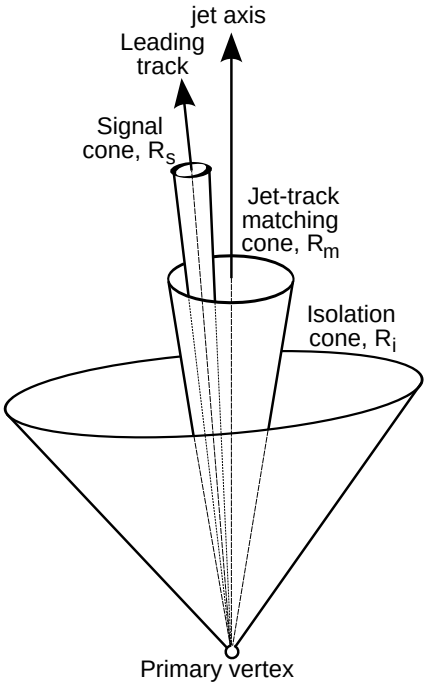

Figure 5.11: Sketch of the basic principle of τ-jet identification with charged track isolation.

In the counting of the tracks, only tracks, whose transverse momentum exceeds a minimum value $\mathrm{p_T^{min}}$ are taken into account. Additionally, tracks, which do not come from the primary vertex, are rejected by requiring that the impact parameter z coordinate of the track matches the impact parameter z coordinate of the leading track within a distance of $\Delta\mathrm{IP_z^{max}}$, i.e. $\Delta\mathrm{IP_z} = |\mathrm{PV_z} - \mathrm{IP_z}| < \Delta\mathrm{IP_z^{max}}$.

The charged track isolation efficiency is shown in Fig. 5.12 for the signal (left plot) and QCD multi-jet background (right plot) for the different $\mathrm{E_T^{MC}}$ bins as a function of the isolation cone size $\mathrm{R_i}$. The algorithm parameters were set to $\mathrm{p_T^{min}} > 1.0$ GeV/c, $\Delta\mathrm{IP_z} <$2.0 mm, $\mathrm{R_m} = 0.1$ to obtain the curves in the figure. The symbols correspond, in the order of decreasing efficiency, to the $\mathrm{E_T^{MC}}$ bins of 130-150, 80-110, 50-70, and 30-50 GeV with filled symbols corresponding to signal cone size of $\mathrm{R_s} = 0.07$ and open symbols corresponding to $\mathrm{R_s} = 0.04$. For good track quality, the tracks were required to contain at least eight hits of which at least two inside the pixel detector. Furthermore, the normalized χ^2 of the track fit was required to be less than 10. The efficiency of Fig. 5.12 was calculated after the MC preselection criteria.

The charged track isolation was observed to provide a rejection factor of $\sim$ 20-50 against hadronic jets from QCD multi-jet events with an efficiency of above 70% for τ jets. Inefficiencies for every step of the charged track isolation algorithm are presented in Table 5.3 for the jets in two bins of $\mathrm{E_T^{MC}}$, 30-50 and 130-150 GeV.

The charged track isolation algorithm is usually followed by a cut on the number of tracks in the signal cone. The hadronic decay products of the τ's consist mainly of one or three charged particles (one and

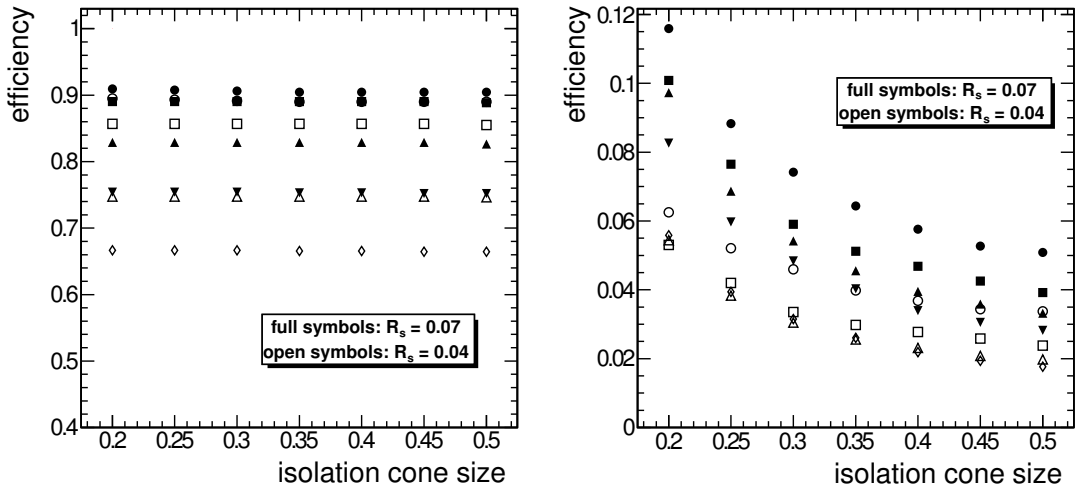

Figure 5.12: Charged track isolation efficiency for the τ jets (left plot) and hadronic jets from QCD multi-jet samples (right plot) as a function of the isolation cone sizes $\mathrm{R_i}$ for the signal cone sizes $\mathrm{R_s}$ = 0.07 (full symbols) and $\mathrm{R_s}$ = 0.04 (open symbols). The symbols correspond in the order of the decreasing efficiency to $\mathrm{E_T^{MC}}$ bins of 130-150, 80-110, 50-70, and 30-50 GeV. The efficiency has been computed after the MC preselection criteria.

Jet $\mathrm{E_T^{MC}}$, GeV	$\geq$1 track in the isol.annulus	Leading track with $\mathrm{p_T} > 6$ GeV/c, $\mathrm{R_m} = 0.1$	Isolation $\mathrm{R_s} = 0.07$	Isolation $\mathrm{R_s} = 0.04$	Total ineff. for $\mathrm{R_s} = 0.04$
30-50	7.7%	10.2%	5.2%	14.2%	32.1%
130-150	4.8%	2.6%	1.0%	2.5%	9.9%

Table 5.3: Inefficiencies for every step of the charged track isolation algorithm for the τ jets in two bins of $\mathrm{E_T^{MC}}$, 30-50 and 130-150 GeV. The isolation annulus is defined as the space between the signal and the isolation cones. The MC preselection have been applied before the charged track isolation.

three-prong, respectively). Therefore, the charged track isolation requirement can be naturally followed by the requirement to have only one or three reconstructed tracks in the signal cone. The one-prong decays represent $\sim$75% of all hadronic decay modes of the τ lepton excluding K^0 decays. The efficiency for the requirement of the number of signal tracks is shown in Table 5.4 for τ jets and hadronic jets from QCD multi-jet events for four bins of $\mathrm{E_T^{MC}}$. The efficiencies in the table were obtained by fixing the isolation and the signal cones to $\mathrm{R_i} = 0.4$ and $\mathrm{R_s} = 0.07$, respectively. To further suppress hadronic jets of low transverse energy, a strong cut can be applied on the $\mathrm{p_T}$ of the leading track. For example, in Ref. [111], a cut of 40 GeV/c was applied on the leading track $\mathrm{p_T}$ to efficiently reduce the QCD multi-jet background in a study of the $\mathrm{H, A} \rightarrow \tau\tau$ channel.

A number of identification methods, which can be applied after the charged track isolation criterion, are discussed in the following subsections. The results presented in these discussions have been preselected with the charged track isolation with the following parameters: $\mathrm{R_m} = 0.1$, $\mathrm{R_s} = 0.07$, $\mathrm{R_i} = 0.4$, $\mathrm{p_T^{min}} > 1$ GeV/c, and $\Delta\mathrm{IP_z} < 2$ mm, unless otherwise stated. One or three tracks were required in the signal cone and the leading track $\mathrm{p_T}$ was required to be greater than 6 GeV/c.

	Jet $\mathrm{E_T^{MC}}$, GeV			
Number of tracks	30-50	50-70	80-110	130-150
Hadronic jets from QCD multi-jet events				
1 track	63%	72%	69%	60%
3 tracks	7%	9%	9%	13%
1 or 3 tracks	70%	81%	78%	73%
τ jets				
1 track	81%	77%	71%	70%
3 tracks	10%	16%	16%	20%
1 or 3 tracks	91%	93%	87%	90%

Table 5.4: Efficiency of the track counting requirement for τ jets and hadronic jets from QCD multi-jet events in four bins of $\mathrm{E_T^{MC}}$.

5.1.7 Identification of τ's with impact parameter

The impact parameter (IP) is the point of closest approach to the beam line. If not otherwise stated, the impact parameter is calculated respective to the primary vertex in this section. The impact parameter is usually used to select tracks, which belong to the primary vertex by applying an upper limit cut on the impact parameter. However, it is also possible to apply a lower limit cut on the impact parameter to take advantage of the relatively long life-time of the τ lepton, which can cause the tracks coming from the τ decay to be offset by a small, but measurable displacement from the beam line.

The unsigned transverse impact parameter ($\mathrm{IP_T}$) distributions of the leading track are shown in Fig. 5.13 for τ jets and for hadronic jets from QCD multi-jet events with one (left plot) and three (right plot) reconstructed tracks in the signal cone. In the one-prong final state, a considerable tail was found to exist for the hadronic jets from QCD multi-jet events for the large transverse impact parameter values (Fig. 5.14). It can be seen from Fig. 5.14 and from the left plot of Fig. 5.13 that this tail can be cut with minor losses on the signal events by placing an upper limit cut at $\mathrm{IP_T} < 300$ μm.

The source of the tails is revealed in Fig. 5.15. Fig. 5.15 shows, that the tails come predominantly from jets, which are emitted in the forward directions. In these jets, the leading track is composed of hits that belong to different simulated tracks in a dense hit environment. By looking at the position of the hits, it is possible to determine where the first change from the original simulated track to another occurs. The left and right plots of Fig. 5.16 show, that the most likely place, where such a 'jump' occurs, is either inside the pixel detector or at the boundary between the pixel and inner strip sensors. Because of the detector layout, the distance between the two consecutive hits where the jump occurs is fairly large at large η. The probability to assign wrong hits to the reconstructed track is thus increased as a result of the long propagation distance in the track building. It was also observed, that the amount of jets in the tail is increased with $\mathrm{E_T^{MC}}$. This behavior is expected, because the track multiplicity increases with $\mathrm{E_T^{MC}}$ and thus also the probability to assign hits from different simulated tracks to the reconstructed track is increased.

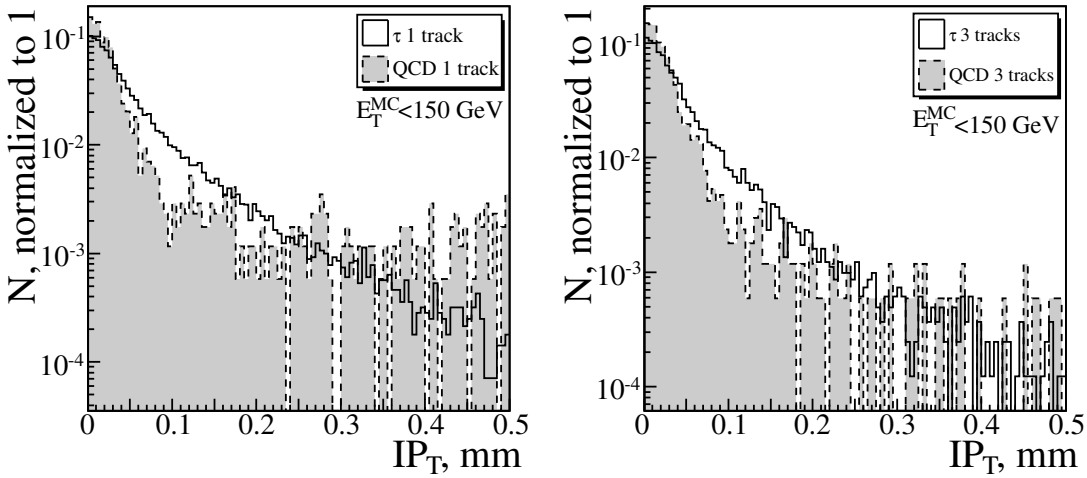

Figure 5.13: Unsigned transverse IP distribution for τ jets and hadronic jets from QCD multi-jet events with one (left plot) or three (right plot) reconstructed tracks. The histograms have been normalized to unity.

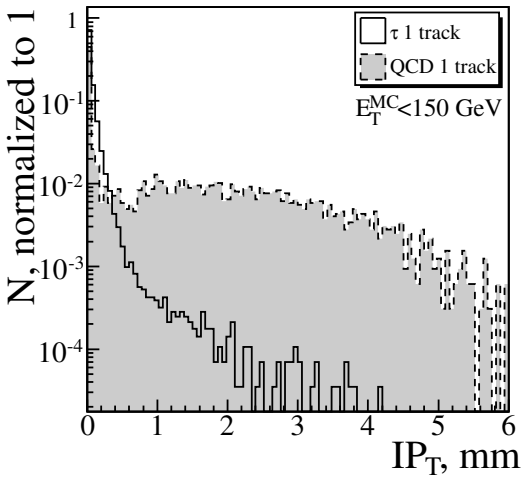

Figure 5.14: Unsigned transverse IP distribution for τ jets and hadronic jets from QCD multi-jet events with one reconstucted tracks and enlarged scale in the x axis. The histograms have been normalized to unity.

It can also be seen in Fig. 5.13, that a lower limit cut on the transverse impact parameter does have rejection power against the hadronic jets from QCD multi-jet events. In order to achieve the best separation power, significance of the impact parameter measurement, i.e. impact parameter value divided by its measurement uncertainty, was used. The measurement uncertainty of the transverse impact parameter was somewhat larger for the background than for the signal events. The mean measurement uncertainty of the transverse impact parameter was found to be 15.0 μm and 16.7 μm for one-prong and three-prong signal events and 17.9 μm and 22.2 μm for one-prong and three-prong background events, respectively. The mean uncertainty on the impact parameter in three-dimensional space was found to be 58.3 μm and 56.6 μm for one-prong and three-prong signal events and 75.7 μm and 60.4 μm for one-prong and three-prong background events, respectively.

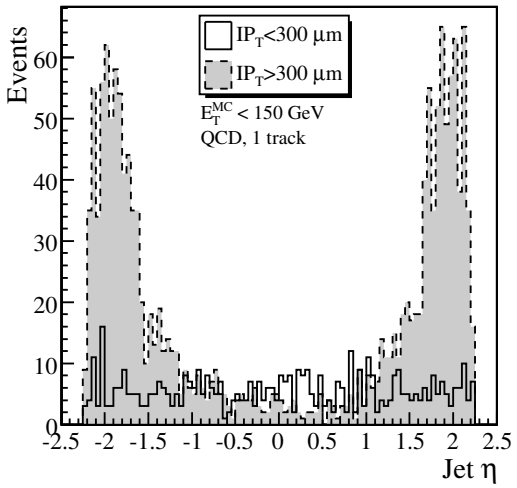

Figure 5.15: Jet η distribution for the hadronic jets from QCD multi-jet events with one reconstructed track for tracks with $\mathrm{IP_T} < 300$ μm and $\mathrm{IP_T} > 300$ μm.

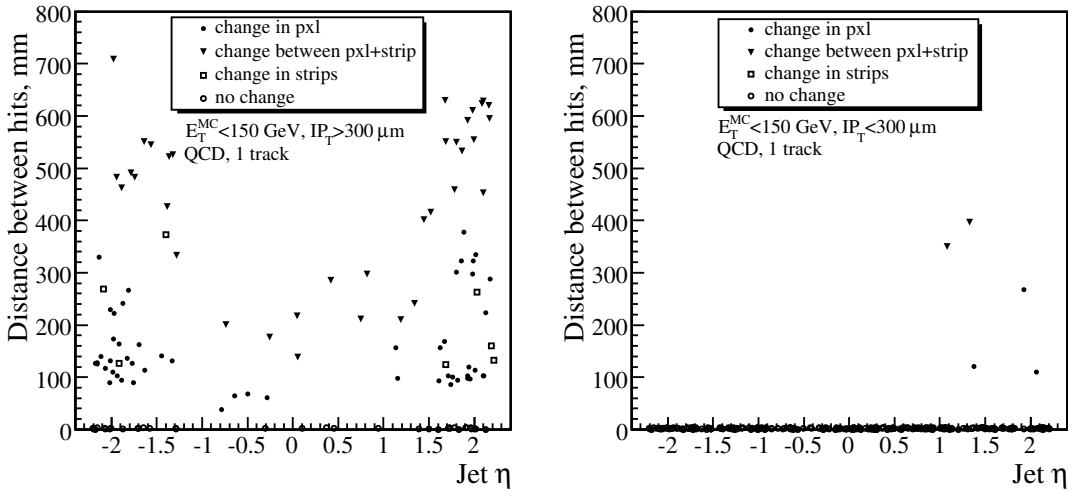

Figure 5.16: The propagation distance between two consecutive reconstructed hits, between which the simulated track association has changed, as a function of the jet η for the $\mathrm{IP_T}$ tail region (left plot) and peak region (right plot) for hadronic jets from QCD multi-jet events.

To determine the performance of the τ-jet identification with impact parameter, the lower limit cut on the impact parameter was varied. The resulting efficiency curves are shown in Fig. 5.17 for jets with one reconstructed track in the signal cone for different bins of jet $\mathrm{E_T^{MC}}$. In the left plot of Fig. 5.17, the efficiency was calculated by varying the lower limit cut on the unsigned impact parameter significance in the transverse plane from 0 to 9 in steps of 1. The right plot of Fig. 5.17 was obtained by varying the lower limit cut on the unsigned impact parameter in three-dimensional space from 1 to 10 in steps of 1. The upper limit cuts of $\mathrm{IP_T} < 300$ μm and $\mathrm{IP} < 1$ mm have been included in the efficiency curves of the left and right plots, respectively, to reject the tails. Since most of the background rejection comes from the rejection of the tails, the efficiency curves of the two plots are very similar. To minimize the effect of the tails, additional cleaning was imposed on the leading track by requiring, that the normalized χ^2

measurement of the fit of the two innermost hits in the leading track was less than 10. The resulting efficiency curves for the transverse and three-dimensional impact parameter significances can be seen in the left and right plots of Fig. 5.18, respectively.

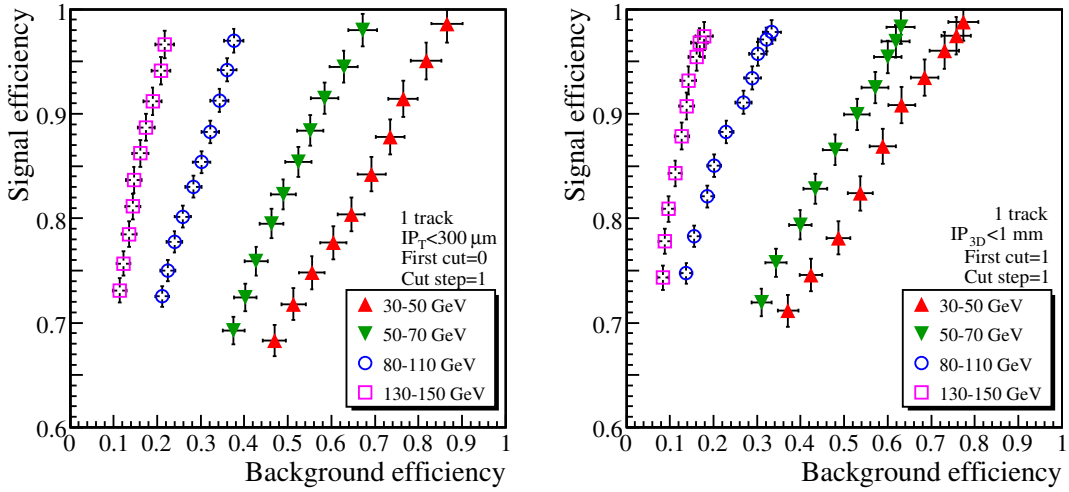

Figure 5.17: Signal and background efficiency curves for four jet $\mathrm{E_T^{MC}}$ bins based on a lower limit cut on the unsigned transverse impact parameter ($\mathrm{IP_T}$) significance (left) and the unsigned three-dimensional IP significance (right) with upper limit cuts of $\mathrm{IP_T} < 300$ μm and $\mathrm{IP} < 1$ mm, respectively. The efficiency has been computed after having applied the MC preselection criteria and the charged track isolation.

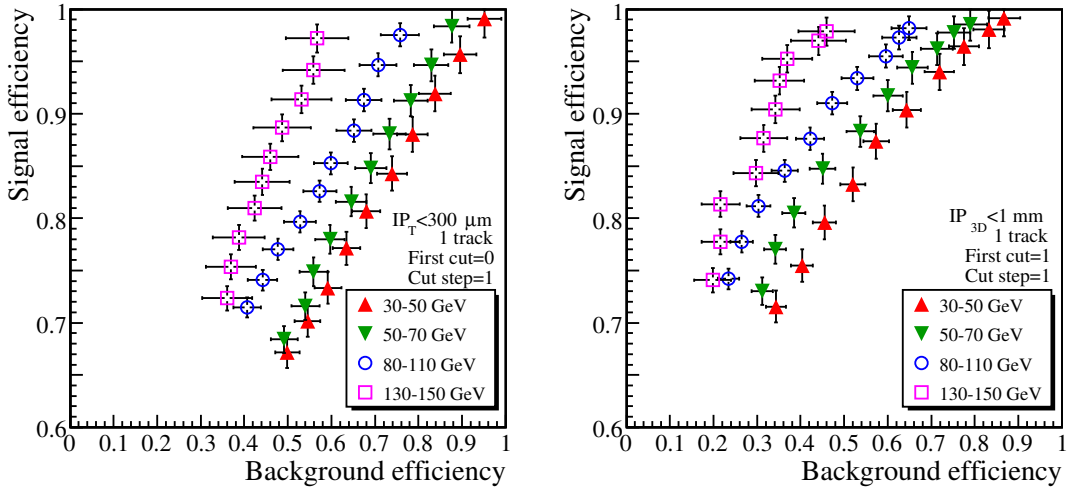

Figure 5.18: Signal and background efficiency curves for different jet $\mathrm{E_T^{MC}}$ bins based on a lower limit cut on the unsigned transverse impact parameter ($\mathrm{IP_T}$) significance (left) and the unsigned three-dimensional IP significance (right). In addition to the track quality cuts, norm. $\chi^2_{\mathrm{hit1}} < 10$ and norm. $\chi^2_{\mathrm{hit2}} < 10$ were required before the efficiency curves. The upper limit cut of $\mathrm{IP_T} < 300$ μm and $\mathrm{IP} < 1$ mm, respectively, is included in the efficiency. The efficiency has been computed after having applied the MC preselection criteria and the charged track isolation.

For jets with three reconstructed tracks in the signal cone, the impact parameter distributions of the leading track were found to be quite similar for the signal and background samples. The obtainable background efficiency was found to be about 81-88% at a signal efficiency of 95% depending on the

jet $\mathrm{E_T^{MC}}$. Attempts were made to combine the impact parameters of the three tracks, but such approach was found to yield only minimal background separation.

To obtain further separation power, a cut on the sign of the impact parameter was studied by applying a minus sign to the impact parameter, if it was reconstructed behind the primary vertex compared to the jet direction. It was, however, observed that the measurement uncertainty caused the impact parameter distributions of the signal were almost symmetrical with respect to the sign of the impact parameter. Since almost half of the signal events would have been cut with this approach, the sign condition was not implemented.

5.1.8 Flight path reconstruction for τ-jet identification

The lifetime of the τ lepton ($c\tau$=87.1$\pm$0.3 μm) allows for the reconstruction of its decay vertex, i.e. secondary vertex (SV), for the three- and five-prong τ decays. Since at least two tracks are required to reconstruct the secondary vertex position, it is not possible to reconstruct the τ flight path in one-prong final states. As a result of the boost effect, the distance, i.e. the flight path, between the primary and secondary vertices is measurable. Since most charged particles in hadronic jets from QCD multi-jet events are produced at the primary vertex, it is possible to acquire background separation by requiring a τ-jet candidate to exceed a certain low cut-off of flight path length.

Table 5.5 shows the fraction of τ jets with a fixed number of reconstructed tracks in the signal cone for MC matched one-, three-, and five-prong τ decays. The proportion of three-prong and five-prong τ decays of τ jets, which have passed the charged track isolation algorithm, was found to be 23.9% and 0.3%, respectively. As the contribution from the five-prong decays is insignificant, only the three-prong τ decays are considered in the following for the flight path study. The case of two reconstructed tracks in the signal cone was not considered, since the QCD multi-jet background was expected to be overwhelming in this final state.

	1 track	2 tracks	3 tracks	$>$ 3 tracks
1-prong τ	88.4$\pm$0.3%	6.1$\pm$0.1%	4.1$\pm$0.1%	1.4$\pm$0.1%
3-prong τ	8.6$\pm$0.1%	16.1$\pm$0.2%	63.2$\pm$0.4%	12.1$\pm$0.2%
5-prong τ	13.1$\pm$1.7%	4.4$\pm$1.0%	11.7$\pm$1.7%	70.9$\pm$4.1%

Table 5.5: The fraction of τ jets with a certain number of reconstructed tracks in the signal cone for one-, three-, and five-prong τ decay.

In the following, the resolution of the secondary vertex and the rejection of background with a cut on the reconstructed τ flight path are discussed.

5.1.8.1 Secondary vertex resolution

The position of the secondary vertex, i.e. the τ lepton decay vertex, was fitted with the Kalman vertex fitter (KVF), adaptive vertex fitter (AVF), and trimmed Kalman vertex fitter (TKF, previously also known as the principal vertex fitter), which have been described in Section 4.3.3.2. The three tracks were used as input to the vertex fitter. To assess the fit of the secondary vertex, the displacement of the reconstructed secondary vertex position from the simulated one ($\mathrm{SV}^{\mathrm{RECO}} - \mathrm{SV}^{\mathrm{MC}}$), i.e. the residual, was calculated. The residual was further divided by the measurement uncertainty of the reconstructed secondary vertex position to obtain the pull values ($\frac{\mathrm{SV}^{\mathrm{RECO}} - \mathrm{SV}^{\mathrm{MC}}}{\sigma^{\mathrm{SV}^{\mathrm{RECO}}}}$).

The secondary vertex residuals and pulls are shown for the signal sample in Figs. 5.19 and 5.20, respectively, for the x and z coordinates and for the different vertex fitters. The residual and pull distributions were fitted with the sum of two Gaussians in order to estimate the center and tail parts of the distributions. The center part of the residual and pull was found to be $\sim$180 μm and $\sim$1.1, respectively, independent of the x and z direction or the vertex fitter used. The tails in the residual plots were found to be coming predominantly from events, in which two of the tracks were very close to each other.

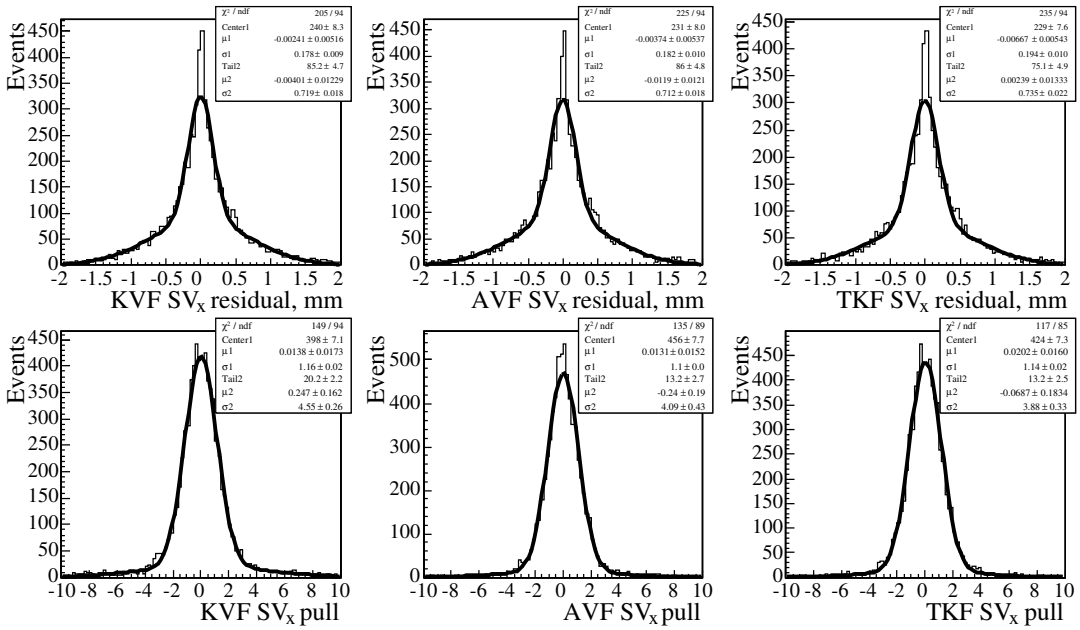

Figure 5.19: The residual (top) and pull (bottom) of the secondary vertex x-coordinate for different vertex fitters. The transverse energy of the jets used for the plot was restricted to $\mathrm{E}_{\mathrm{T}}^{\mathrm{MC}} < 150$ GeV.

The resolution of the secondary vertex fit is considerably better in the plane transverse to the jet axis than in the direction along the jet axis, because the three tracks are very collimated because of the boost effect. Hence, it is more natural to operate in a jet coordinate system instead of the global CMS detector coordinate system. The transition from the global coordinate system ($\vec{\mathrm{x}}, \vec{\mathrm{y}}, \vec{\mathrm{z}}$) to the jet coordinate system

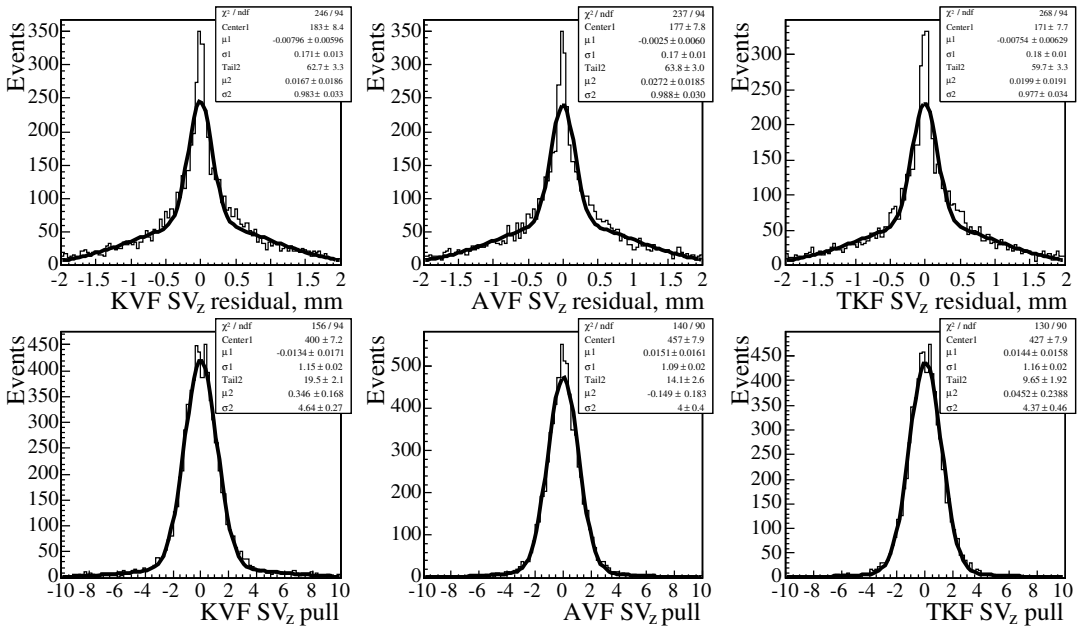

Figure 5.20: The residual (top) and pull (bottom) of the secondary vertex z-coordinate for different vertex fitters. The transverse energy of the jets used for the plot was restricted to $\mathrm{E_T^{MC}} < 150$ GeV.

$(\vec{\mathrm{i}}, \vec{\mathrm{j}}, \vec{\mathrm{k}})$ can be defined as

$$\begin{cases} \vec{\mathrm{k}} &= \text{jet axis direction} \\ \vec{\mathrm{i}} &= \vec{\mathrm{k}} \times \vec{z} \\ \vec{\mathrm{j}} &= \vec{\mathrm{k}} \times \vec{\mathrm{i}} \end{cases} . \tag{5.2}$$

With this definition, the residual vector of the global coordinate system can be projected to the jet coordinate system while preserving the sign of each component.

The resolution of the secondary vertex reconstruction is shown in Fig. 5.21 in the plane transverse to the τ-jet axis (left plot) and in the direction along the τ-jet axis as a function of the jet $\mathrm{E_T^{MC}}$. In the plane transverse to the τ-jet axis, the resolution $\sigma_{\mathrm{transverse}}$ of the reconstructed secondary vertex was observed to be between 18 and 25 μm independent of the τ-jet $\mathrm{E_T^{MC}}$, when $\mathrm{E_T^{MC}}$ was varied between 30 and 300 GeV. The resolution in the direction parallel to the τ-jet axis ($\sigma_{\mathrm{long.}}$) was found to be $\sim$600 μm for τ-jet $\mathrm{E_T^{MC}}$ of 30 GeV and to increase with $\mathrm{E_T^{MC}}$. The three different vertex fitters were observed to deliver almost equal performance for the resolution in both the plane transverse to the τ-jet direction and the direction along the τ-jet axis.

5.1.8.2 Identification τ jets with τ flight path reconstruction

The flight path vector was calculated by substracting the position of the secondary vertex from the position of the primary vertex. The flight path was assigned a minus sign, if the secondary vertex was reconstructed behind the primary vertex compared to the jet direction. The measurement uncertainty of the flight path

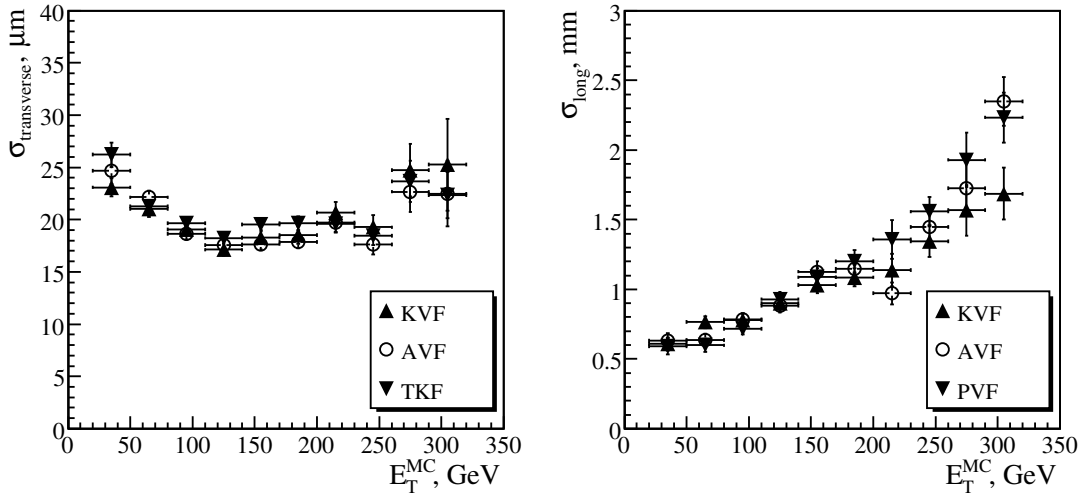

Figure 5.21: The resolution of the central Gaussian of the reconstructed secondary vertex projected transverse to the τ jet axis (left) and parallel to τ jet direction (right) for the different vertex fitters as a function of τ-jet $\mathrm{E_T^{MC}}$.

is given in matrix notation by

$$\sigma = \sqrt{\frac{\mathrm{LEL^T}}{\mathrm{LL^T}}}, \tag{5.3}$$

where L is the flight path column vector and E is the sum of the covariance matrices of the primary and secondary vertex measurements. The reconstruction of the primary vertex is described in Section 5.1.1. It was observed, that the uncertainty in the flight path reconstruction was completely dominated by the uncertainty of the secondary vertex.

The reconstructed transverse flight path of the three-prong τ-jet candidates after charged track isolation is shown in Fig. 5.22 for signal τ jets (left plot) and hadronic jets from QCD multi-jet events (right plot). The reconstruction of the secondary vertex for the τ jets poses a challenge, because the three tracks in the signal cone are very collimated in the pixel layers because of the boost effect. Because of the pixel sensor granularity, it is possible that hits in the pixel sensors are shared between different tracks, which can bias the position of the secondary vertex. The fraction of τ jets with at least two shared hits in the reconstructed tracks was found to be 9.6% and 22.8% for τ jets with $\mathrm{E_T^{MC}} < 150$ GeV and $150 < \mathrm{E_T^{MC}} < 420$ GeV, respectively, compared to the total number of three-prong τ jets with the respective energy after charged track isolation. The bias caused by the shared hits is visible in Fig. 5.22 as bumps at $\sim$40 mm and $\sim$70 mm from the beam axis, which correspond to the location of the first and the second barrel layers of the Pixel detector. The amount of biased jets can be observed to increase with τ jet $\mathrm{E_T^{MC}}$ in the left plot of Fig. 5.22. This effect is expected, since the boost not only causes the tracks to be more collimated as a function of τ jet $\mathrm{E_T^{MC}}$, but it also increases the average flight path length of the τ lepton hence leaving the tracks less space to curve before traversing the first pixel detector layer.

The b and c jets can pose a potential hazard for τ identification with flight path reconstruction, since their decay chains involve the production of b- and c-hadrons, which have a flight path comparable to that

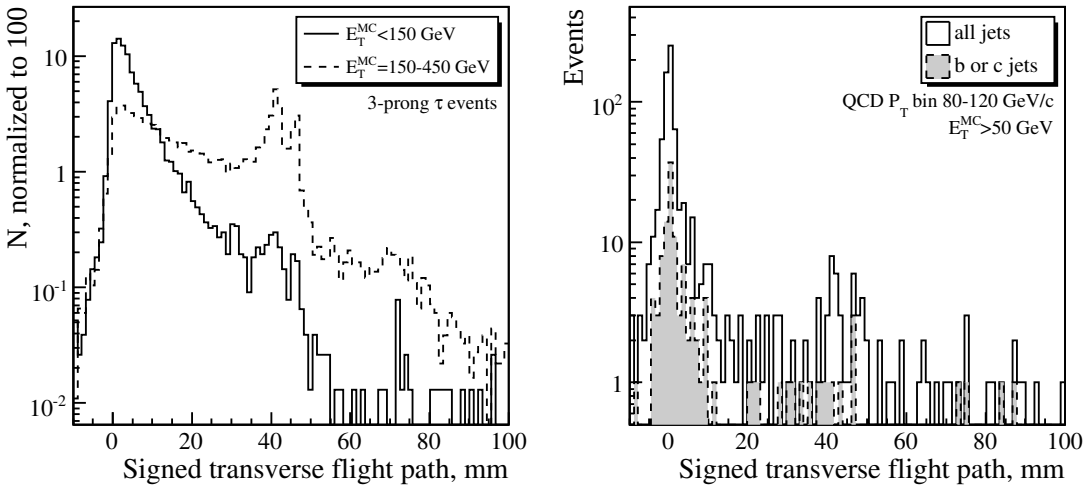

Figure 5.22: Signed transverse flight path for τ jets (left) and hadronic jets from QCD multi-jet events (right), which have passed the charged track isolation algorithm with three reconstructed tracks in the signal cone. The detailed explanation can be found in the text. The histograms have been normalized to 100.

of τ leptons. It was, however, observed, that the charged track isolation algorithm and the requirement to have only three charged tracks in the signal cone of the charged track isolation algorithm effectively suppresses such jets. As can be seen in the right plot of Fig. 5.22, the b- and c-jets do not pose a serious risk of signal contamination. The fraction of b- and c-jets after the charged track isolation is shown in Table 5.6 for different intervals of $\mathrm{E_T^{MC}}$ compared to all hadronic jets from QCD multi-jet events in the given $\mathrm{E_T^{MC}}$ interval. The effect resulting from the hit sharing was found to be observable also for the hadronic jets from QCD multi-jet events as can be seen from the slight bump at $\sim$40 mm in the right plot of Fig. 5.22.

Jet $\mathrm{E_T^{MC}}$ bin, GeV	Fraction of c jets, %	Fraction of b jets, %
30-50	13.1$\pm$2.1	3.1$\pm$1.0
50-70	12.3$\pm$1.6	4.4$\pm$1.0
80-110	13.4$\pm$1.4	2.7$\pm$0.6
130-150	12.4$\pm$1.9	3.0$\pm$1.0

Table 5.6: Fraction of the c- and b-jets in hadronic jets from QCD multi-jet events in four bins of $\mathrm{E_T^{MC}}$ after having passed charged track isolation with three reconstructed tracks in the signal cone.

The performance of the τ-jet identification based on a cut on the reconstructed flight path is shown in Fig. 5.23 for the signed transverse flight path length (left plot) and significance (right plot) for the different jet $\mathrm{E_T^{MC}}$ bins after charged track isolation. The points in the left plot correspond in decreasing order of signal significance to the cut on the signed flight path length varied between 0 to 4.5 mm in steps of 0.5 mm. In the right plot, the points correspond in decreasing order of signal significance to the cut on the signed flight path significance varied between 0 to 4.5 in steps of 0.5. The Kalman vertex fitter described in Section 4.3.3.2 was used to fit the secondary vertex. In order to remove the bias from the

sharing of hits, the transverse flight path value was required to be less than 35 mm, which is well below the first layer of pixel detectors. It can be concluded from the right plot of Fig. 5.23 that a rejection factor of ∼5 can be achieved with a signal efficiency of 70-80% for τ jets with $\mathrm{E_T^{MC}}$ between 30 and 170 GeV.

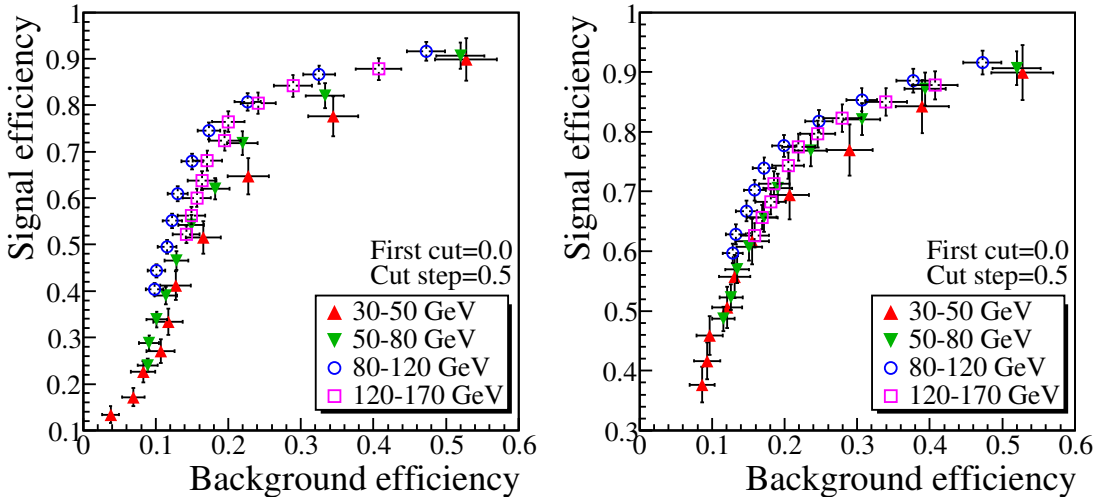

Figure 5.23: Performance of τ-jet identification with a cut on the reconstructed signed transverse τ flight path (left) and on the signed transverse τ flight path significance (right) for the different jet $\mathrm{E_T^{MC}}$ bins. The efficiency has been computed after applying MC preselection criteria and the charged track isolation.

5.1.9 Reconstruction of the visible τ-jet mass

The visible τ-jet mass is a distinct peak below m_{τ} = 1.78 GeV/c^2 smeared and shifted somewhat because of the presence of the τ neutrino, whereas the invariant mass of many hadrons occurring in hadronic jets from QCD multi-jet events are heavier than τ leptons and produce a broader invariant mass spectrum. To exploit this property, the visible τ mass, i.e. τ mass excluding the momentum of the neutrino, is reconstructed from the momentum of the tracks in the signal cone of the charged track isolation algorithm and from the energy of the clusters in the electromagnetic calorimeter within a certain cone in η,ϕ space around the τ-jet axis. Since both the charged and neutral energy is used for the reconstruction, the method can be applied to both one- and three-prong final states of τ decays.

The dependence of the transverse energy of ECAL clusters, $\mathrm{E_T^{em.}}$, on the distance $\Delta\mathrm{R_{jet}}$ between the jet axis and the ECAL clusters position is shown in Fig. 5.24 for τ jets (left plot) and hadronic jets from QCD multi-jet events (right plot) with $\mathrm{E_T^{MC}}$ between 30 and 150 GeV. It can be seen in the figure, that the $\Delta\mathrm{R_{jet}}$ and $\mathrm{E_T^{em.}}$ are strongly correlated and hence a constraint on one variable does naturally restrict the other one. The ECAL clusters were found to be closer to the jet axis for the τ jets than for the hadronic jets from QCD multi-jet events. For both the τ jets and the hadronic jets from QCD multi-jet events, the ECAL clusters with $\mathrm{E_T^{em.}} > 10$ GeV were can be seen to be contained within the cone of the size 0.1 around the calorimeter jet axis.

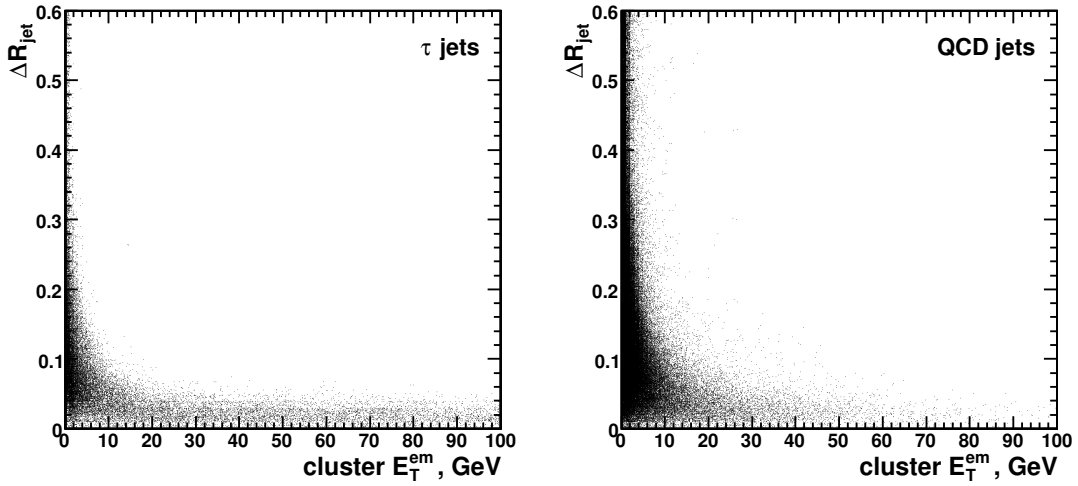

Figure 5.24: Dependence of the transverse energy of the ECAL clusters, $\mathrm{E_T^{em.}}$, and the distance $\Delta\mathrm{R_{jet}}$ in the η,ϕ space between the calorimeter jet axis and the ECAL clusters for the τ jets (left plot) and hadronic jets from QCD multi-jet events for $\mathrm{E_T^{MC}}$ between 30 and 150 GeV (right plot) after charged track isolation.

The ECAL energy was summed in a cone of 0.40 around the jet axis direction. The cone size of 0.4 was found to be the optimal for the τ-jet identification performance. A smaller cone size was observed to reduce the background separation. When summing the transverse energy of the ECAL clusters, care had to be taken in order not to count the energy left in the ECAL by charged particles. To avoid such double counting of the energy, the ECAL clusters were matched to tracks. Only ECAL clusters, which were separated from track impact points on the ECAL surface by the distance $\Delta\mathrm{R_{track}} > 0.08$ in η,ϕ space, were taken into account. The resulting reconstructed invariant τ mass $\mathrm{M_{\tau}^{jet}}$ is shown for τ jets in Fig. 5.25 with (dashed line) and without (solid line) double counting of the charged energy. It can be seen in the figure, that the tail is reduced, if double-counting is disabled. The peak at zero value of $\mathrm{M_{\tau}^{jet}}$ was found to be caused by τ jets with a single track and no ECAL clusters satisfying the constraints on $\mathrm{M_{\tau}^{jet}}$.

The invariant mass distributions for different $\mathrm{E_T^{MC}}$ bins of τ jets and hadronic jets from QCD multi-jet events are shown in Fig. 5.26. The corresponding selection efficiency of the cut $\mathrm{M_{\tau}^{jet}} < 2.5$ GeV/c^2 is shown in Table 5.7 for τ jets and hadronic jets from QCD multi-jet events. The efficiency for the τ jets was found to depend only slightly on the jet $\mathrm{E_T^{MC}}$, whereas the rejection factor against for hadronic jets from QCD multi-jet events was observed to strongly increase with the $\mathrm{E_T^{MC}}$ of the jets.

The visible τ-jet mass can also be calculated from the four-momenta of the charged tracks. Although such an approach is very robust and does not have the problem of potential double counting of charged energy, its drawback is that it is only applicable to the final state of three charged tracks.

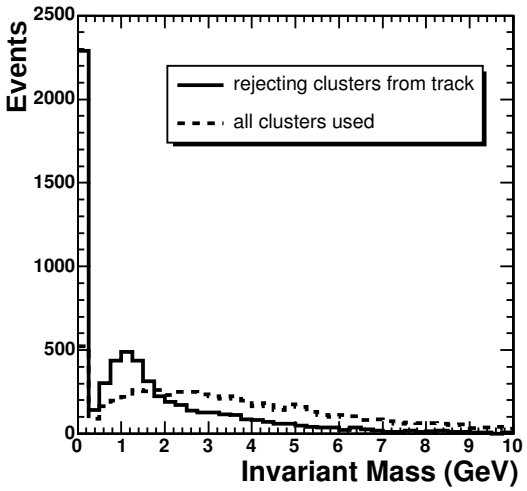

Figure 5.25: The invariant mass distribution for τ jets when all ECAL clusters were taken in a cone of 0.4 around the jet axis (dashed line) and when ECAL clusters not matched to tracks were used (solid line).

	Jet $\mathrm{E_T^{MC}}$, GeV			
	30-50	50-70	80-110	130-150
Efficiency for τ jets,%	86.32	82.27	83.02	80.76
Efficiency for hadronic jets,%	33.67	19.16	6.05	2.47

Table 5.7: Selection efficiency of the cut $\mathrm{M_{\tau}^{jet}} < 2.5$ GeV/c^2 for the τ jets and the hadronic jets from QCD multi-jet events for different intervals of $\mathrm{E_T^{MC}}$. The efficiency has been computed after having applied the MC preselection criteria and the charged track isolation.

5.1.10 Rejection of electrons and muons

A genuine electron can pass all the τ-jet identification criteria described so far. To suppress the rate of electrons, which are misidentified as τ jets in the offline analysis, two methods of similar performance were tested. In the first method, the electron rejection is done with a lower cut-off on the transverse energy of the maximal HCAL tower belonging to the reconstructed jet. In the second method a lower cut-off was applied on the value of $\mathrm{E_T^{hadr.}}/\mathrm{p_T^{ltr}}$, where $\mathrm{E_T^{hadr.}}$ is the transverse energy of the jet measured with HCAL alone and $\mathrm{p_T^{ltr}}$ is the transverse momentum of the leading track.

In the following, the performance of the first method is presented. The transverse energy of the HCAL cell with the highest transverse energy is shown in Fig. 5.27 for an electron with $\mathrm{p_T} = 35$ GeV/c reconstructed as a jet (solid line) and for τ jets with true τ-jet transverse energy between 40-60 GeV (dashed line) and 100-140 GeV (dotted line). The transverse momentum of the leading track was required to exceed 10 GeV/c. It can be seen in the figure, that for most electrons, the energy of the HCAL cell was rather small as expected. The high $\mathrm{E_T}$ tail for the true electrons in Fig. 5.27 was found to correspond to electrons going through the η/ϕ cracks in the ECAL modules, and gaps between the barrel and endcap

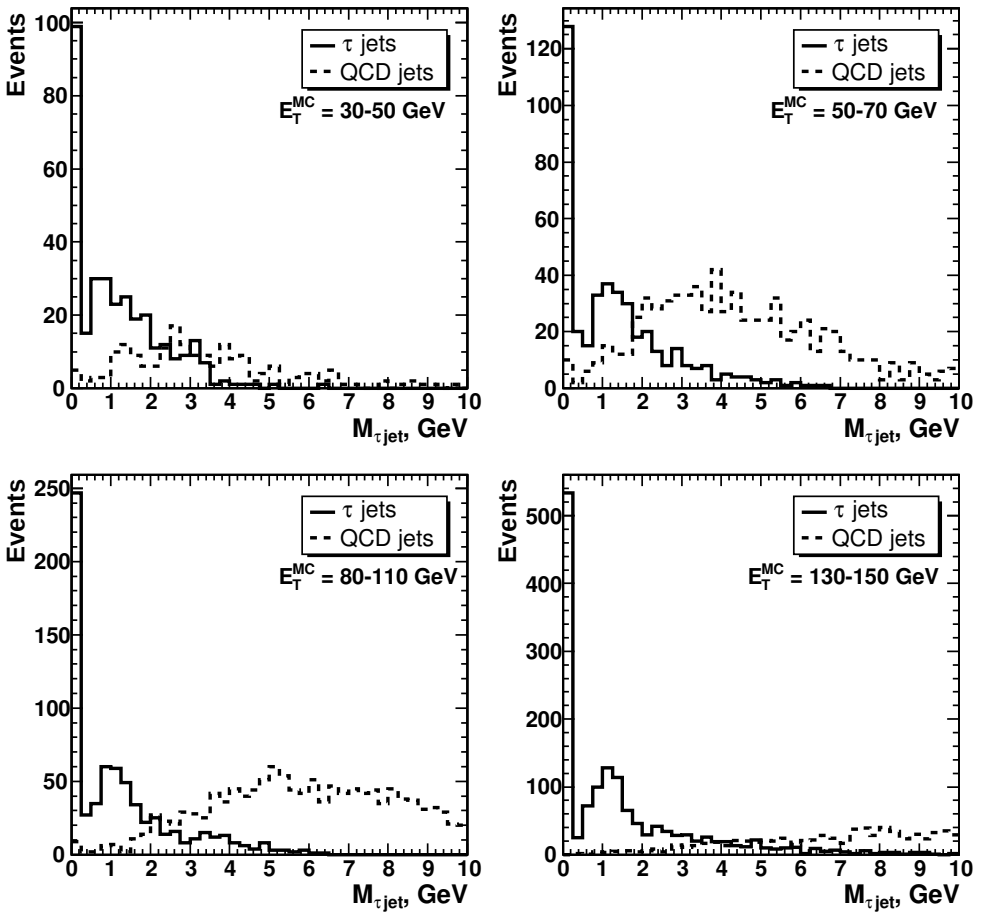

Figure 5.26: Distribution of $\mathrm{M}_{\tau}^{\mathrm{jet}}$ of τ jets (solid line) and hadronic jets from QCD multi-jet events (dashed line) in the $\mathrm{E}_{\mathrm{T}}^{\mathrm{MC}}$ bins of 30-50 GeV (top-left plot), 50-70 GeV (top-right plot), 80-110 GeV (bottom-left plot) and 130-150 GeV (bottom-right plot).

detectors. The signal efficiency and the corresponding electron contamination are shown in Table 5.8 for the requirement of the HCAL cell with the highest transverse energy to exceed either 1 GeV or 2 GeV.

Cut	Electron contamination	τ jet $\mathrm{E_T}$ 40-60 GeV		τ jet $\mathrm{E_T}$ 100-140 GeV	
		$\mathrm{p}_{\mathrm{T}}^{\mathrm{ltr}} > 10$ GeV/c	$\mathrm{p}_{\mathrm{T}}^{\mathrm{ltr}} > 25$ GeV/c	$\mathrm{p}_{\mathrm{T}}^{\mathrm{ltr}} > 10$ GeV/c	$\mathrm{p}_{\mathrm{T}}^{\mathrm{ltr}} > 25$ GeV/c
$>$1 GeV	0.08	0.936	0.971	0.977	0.991
$>$2 GeV	0.03	0.854	0.917	0.942	0.969

Table 5.8: Efficiency of the cut on the transverse energy of the maximal $\mathrm{E_T}$ HCAL tower for a jet reconstructed from an electron with $\mathrm{p_T}$ = 35 GeV/c and for a τ jet in two $\mathrm{E}_{\mathrm{T}}^{\mathrm{MC}}$ ranges of the and for two cuts on transverse momentum of the leading track ($\mathrm{p}_{\mathrm{T}}^{\mathrm{ltr}}$) of the τ-jet candidate.

The misidentification of a muon as a τ jet was not considered, since the average energy losses of a muon in the calorimeter are of the order of a few GeV, which is well below the lowest $\mathrm{E_T}$ threshold for a τ jet to be used in the analysis of the $\mathrm{H, A} \rightarrow \tau\tau$ and $\mathrm{H}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decay channels.

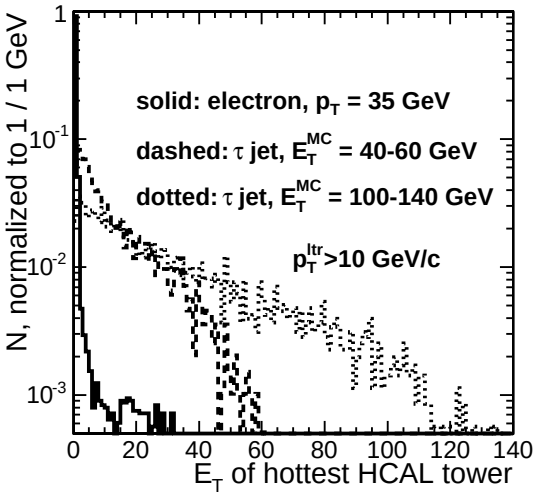

Figure 5.27: Transverse energy of the maximal $\mathrm{E_T}$ HCAL tower belonging to the reconstructed jet for electrons of $\mathrm{p_T}$ = 35 GeV/c (solid line), and τ jets with $40 < \mathrm{E_T^{MC}} < 60$ GeV (dashed line) and $100 < \mathrm{E_T^{MC}} < 140$ GeV (dotted line). The transverse momentum of the leading track was required to exceed 10 GeV/c. All histograms have been normalized to unity.

5.1.11 Rejection of neutral hadrons

In three-prong τ decays it can be of advantage to select only the three-prong final states without π^0's, which constitute about 2/3 of all three-prong final states. In such cases, the energy carried by the charged tracks should match well the visible jet energy. Hence, a cut can be placed on the variable

$$\Delta\mathrm{E} = \frac{\sum \mathrm{E_{tracks}}}{\mathrm{E_{jet}}} - 1 \tag{5.4}$$

by requiring $\Delta\mathrm{E}$ to be contained within a symmetrical window around the value zero. The rejection power of this cut comes from the fact, that hadronic jets from QCD multi-jet events tend to carry larger fraction of their energy in π^0's and in particular neutral hadrons compared to τ jets [25].

5.1.12 Helicity correlations

In some physics processes, it is possible to separate τ leptons of the signal process from τ leptons produced in background processes because of the different spin-parity correlations of the τ lepton production processes. This is especially interesting for separating τ leptons coming from the $\mathrm{H}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decay from τ's which are coming from the $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decay.

To assess how the simulated data samples correspond to the phenomenology of helicity correlations described in Section 2.3.2, event samples described in Section 7.7.1.1 were used. The $\tau^{\pm} \rightarrow \rho^{\pm}\nu_{\tau}$ decay mode was used to determine the fraction of transversally and longitudinally polarized vector mesons in the decay of τ leptons coming from the $\mathrm{H}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decay as a function of jet $\mathrm{E_T}$. In Section 2.3.2.1 it was

established, that the $\rho_{\mathrm{T}}^{\pm}$ decays produce an $\mathrm{R}_{\tau} = \mathrm{p}^{\mathrm{ltr}}/\mathrm{E}^{\mathrm{vis.}\ \tau\ \mathrm{jet}}$ distribution peaked around $\mathrm{R}_{\tau} = 0.5$ because of even energy distribution of the vector meson amongst the pions, whereas the $\rho_{\mathrm{L}}^{\pm}$ decays produce on an R_{τ} distribution peaked at $\mathrm{R}_{\tau} = 0.2$ or $\mathrm{R}_{\tau} = 0.8$ because of uneven energy distribution of the vector meson amongst the pions. Therefore, the ratio of τ jets with R_{τ} within a window around $\mathrm{R}_{\tau} = 0.5$ compared to all τ yields an estimate how much of the transversally and longitudinally polarized vector mesons are left in the event sample after a cut on jet E_{T} threshold.

It can be observed in Fig. 5.28, that for signal τ jets matched to the MC τ lepton from the $\mathrm{H}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decay in the $\tau^{\pm} \rightarrow \rho^{\pm}\nu_{\tau}$ final state, both $\rho_{\mathrm{L}}^{\pm}$ and $\rho_{\mathrm{T}}^{\pm}$ polarization modes are present, for τ jets with $\mathrm{E}_{\mathrm{T}} < 40$ GeV. As the jet E_{T} is increased, it can be seen in the figure, that the longitudinally polarized vector meson decays start to dominate. It can be established from Fig. 5.28, that for $\mathrm{H}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ events with $\mathrm{m}_{\mathrm{H}^{\pm}}$ = 217-409 GeV/c^2 (m_{A} = 200-400 GeV/c^2), a saturation point is reached at jet E_{T} = 80 GeV, beyond which the vector mesons coming from the signal τ decays are predominantly longitudinally polarized. It can also be seen in Fig. 5.28, that the saturation point for the ratio of transversally or longitudinally polarized vector mesons is reached at lower jet E_{T} values, if the $\mathrm{m}_{\mathrm{H}^{\pm}}$ is decreased. Therefore, the τ jets coming from the $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decays must have predominantly transversally polarized vector mesons at jet E_{T} values above 80 GeV, because $\mathrm{m}_{\mathrm{H}^{\pm}} > \mathrm{m}_{\mathrm{W}^{\pm}}$.

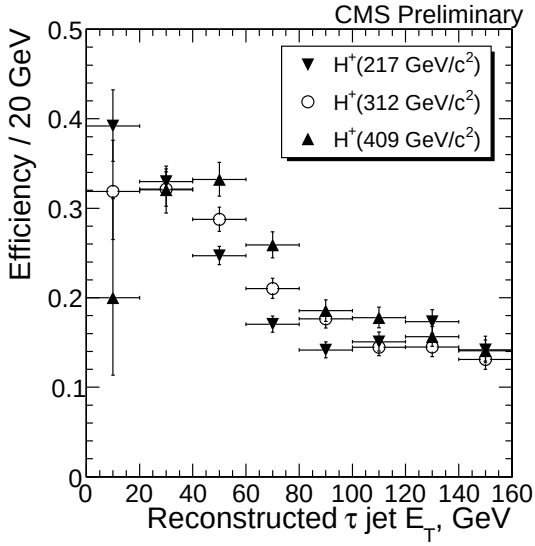

Figure 5.28: Fraction of MC matched τ jets with the $\tau^{\pm} \rightarrow \rho^{\pm}\nu_{\tau} \rightarrow \pi^{\pm}\pi^{0}\nu_{\tau}$ final states passing a cut of $0.35 < \mathrm{R}_{\tau} < 0.65$ as a function of the reconstructed τ jet E_{T} for the signal samples without other cuts.

The R_{τ} distributions of the MC matched signal τ jets and the τ jets from $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decay shown in Fig. 5.29 can be seen to agree well with the phenomenological R_{τ} distributions in the different one-prong τ decay final states with jet $\mathrm{E}_{\mathrm{T}} > 119$ GeV and after having passed charged track and electromagnetic isolation. The R_{τ} distributions are smeared in Fig. 5.29 as a result of contamination of other particles from the underlying event and because of the uncertainty of the jet energy scale. The most clean decay mode, $\tau^{\pm} \rightarrow \pi^{\pm}\nu_{\tau}$, shown in Fig. 5.29 provides an opportunity to estimate the spread of the distributions

resulting from these effects.

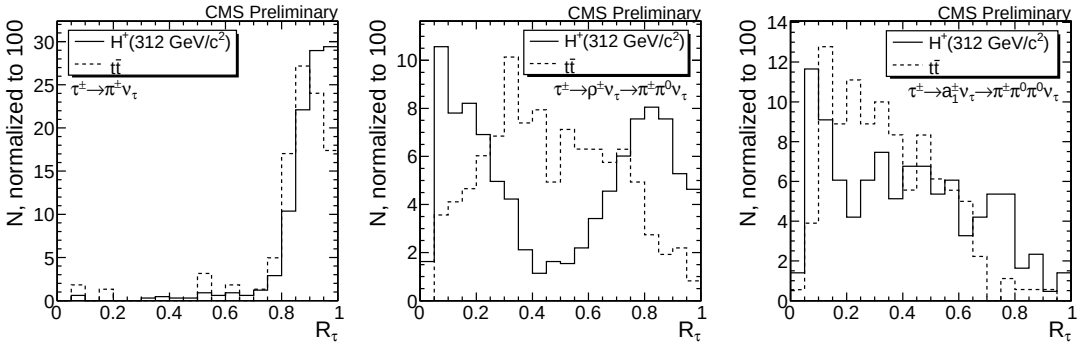

Figure 5.29: Distribution of the R_{τ} variable for MC matched τ jets of the signal (solid) and $\mathrm{t\bar{t}}$ (dashed) samples in the decay mode $\tau \rightarrow \pi^{\pm}\nu_{\tau}$ (left), $\tau \rightarrow \rho^{\pm}\nu_{\tau} \rightarrow \pi^{\pm}\pi^{0}\nu_{\tau}$ (middle), and $\tau \rightarrow \mathrm{a}_{1}^{\pm}\nu_{\tau} \rightarrow \pi^{\pm}\pi^{0}\pi^{0}\nu_{\tau}$ (right) after all other cuts except for the cut on the $\mathrm{p_T}$ of the track.

When the final states of the one-prong τ decay modes are combined together (Fig. 5.30), it is evident, that the τ's coming from the $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decays can be suppressed by requiring the charged pion to carry the majority of the visible τ jet energy. Typically, the R_{τ} variable of the τ jet candidates is required to exceed $\sim$0.7-0.8. The nominal single τ trigger requirement of the leading track $\mathrm{p_T}$ to exceed 20 GeV/c was applied in the figure and it can be seen to suppress the signal strongly in the case where the charged pion carried very little of the visible τ jet energy.

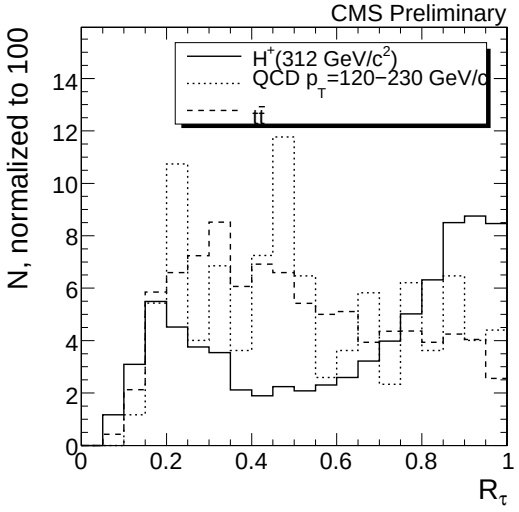

Figure 5.30: Distribution of the R_{τ} variable for MC matched τ jets of the signal (solid) and $\mathrm{t\bar{t}}$ (dashed) samples and for τ-jet candidates from the QCD multi-jet events with $\hat{\mathrm{p}}_{\mathrm{T}}$ = 120-230 GeV/c (dotted) for the one-prong final state after the cut on the jet $\mathrm{E_T}$, leading track $\mathrm{p_T}$, electromagnetic isolation, charged track isolation, and electron rejection.

The R_{τ} cut can be seen to suppress efficiently also the hadronic jets from QCD multi-jet events in Fig. 5.30. The main reason for this effect is, that the hadronization process of the jets tends to produce neutral particles and that there are no helicity correlations in hadronic jets. Therefore, it is quite rare, that a single charged pion does carry the majority of the visible jet energy in the hadronic jets assuming, that the

visible jet energy can be determined with reasonable accuracy. The choice of the fragmentation function can obviously change the number of hadronic jets for which $\mathrm{R}_{\tau} \sim 1$.

For the three-prong τ decay without neutral pions, $\tau \rightarrow \mathrm{a}^{\pm}_{1\mathrm{L,T}}\nu_{\tau} \rightarrow \pi^{\pm}\pi^{\pm}\pi^{\mp}\nu_{\tau}$, the ratio of the momentum carried by the same signed pions against the visible τ jet energy, $\mathrm{R}_{\tau 3} = \mathrm{p}_{\pi^{\pm}\pi^{\pm}}/\mathrm{E}$, has been suggested to be used in Ref. [46] (left plot of Fig.5.31). Better performance is, however, obtained, if the momentum of the leading track is compared to the visible τ-jet energy, i.e. $\mathrm{R}_{\tau} = \mathrm{p}_{\mathrm{ltr}}/\mathrm{E}_{\mathrm{vis.\ \tau\ jet}}$. The R_{τ} variable was found to give a better separation against the $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decay, as can be seen from the following reasoning. For the signal τ decays, where the two same signed pions carry most of the visible τ jet energy, the leading track is usually one of the two same signed pions. On the other hand, if the two same signed pions carry little of the visible jet energy, the leading track is usually the opposite signed pion. In both cases, the leading track carries more energy than the two other pions and thus the R_{τ} value is considerably higher than 0.33, with a substantial tail reaching to higher R_{τ} values (right plot of Fig. 5.31). It was observed, that if the two same signed pions carry almost as much of the a_1 energy as the oppositely signed pion, the R_{τ} peaks at $\mathrm{R}_{\tau} \sim 0.33$. For the $\mathrm{a}^{\pm}_{1\mathrm{T}} \rightarrow \pi^{\pm}\pi^{\pm}\pi^{\mp}$ decay, which is dominant for the τ's from the $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decay, the expected average R_{τ} value is $\mathrm{R}_{\tau} \sim 0.5$ with a decreasing tail to the higher R_{τ} values. It can be seen in Fig. 5.32, that placing a threshold on $\mathrm{R}_{\tau} \sim 0.5$ or greater yields separation of background from signal. Also the separation of signal from hadronic jets was found to be better for the R_{τ} variable than for the $\mathrm{R}_{\tau 3}$ variable.

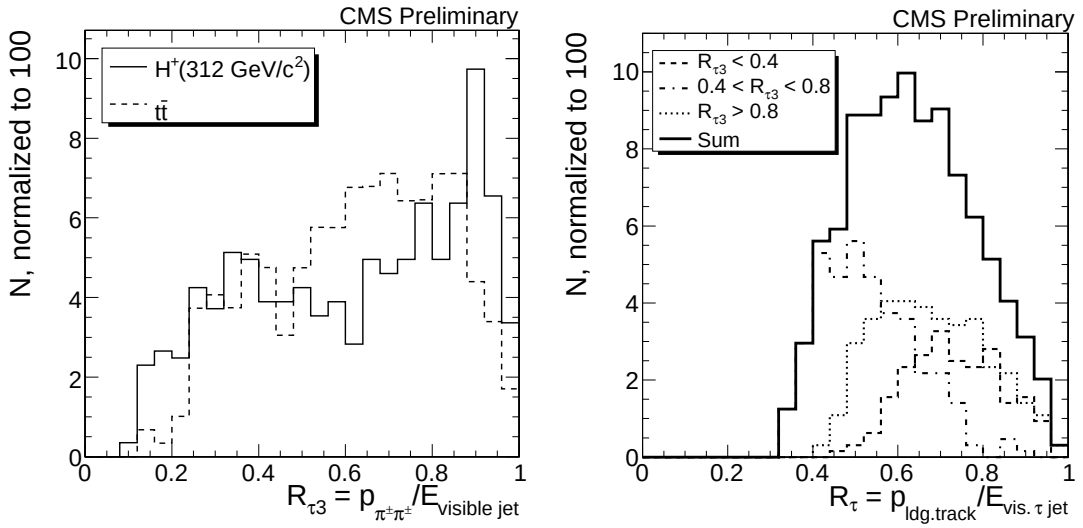

Figure 5.31: Left: the MC matched $\mathrm{R}_{\tau 3}$ distribution for the $\mathrm{m}_{\mathrm{H}^{\pm}}$ =312 GeV/c^2 signal sample (solid) and the $\mathrm{t\bar{t}}$ sample (dashed) in the $\tau^{\pm} \rightarrow \pi^{\pm}\pi^{\pm}\pi^{\mp}\nu_{\tau}$ decay mode. Right: the MC matched R_{τ} distribution for the $\mathrm{m}_{\mathrm{H}^{\pm}}$ =312 GeV/c^2 signal sample (solid) and the components $\mathrm{R}_{\tau 3} < 0.4$ (dashed), $0.4 < \mathrm{R}_{\tau 3} < 0.8$ (dotted), and $\mathrm{R}_{\tau 3} > 0.8$ (dash-dotted) in the $\tau^{\pm} \rightarrow \pi^{\pm}\pi^{\pm}\pi^{\mp}\nu_{\tau}$ decay mode. In both plots, the jet has been required to have $\mathrm{E}_{\mathrm{T}} > 119$ GeV, leading track is asked to have $\mathrm{p}_{\mathrm{T}} > 20$ GeV/c, and the charged track and the electromagnetic isolation have been applied.

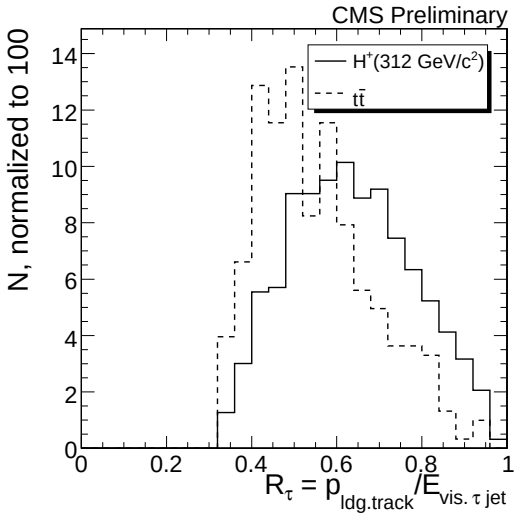

Figure 5.32: The MC matched R_{τ} distribution for the $\mathrm{m_{H^{\pm}}}$ =312 GeV/c^2 signal sample (solid) and the $\mathrm{t\bar{t}}$ sample (dashed) in the $\tau^{\pm} \rightarrow \pi^{\pm}\pi^{\pm}\pi^{\mp}$ decay mode after the cut on the jet $\mathrm{E_T}$, leading track $\mathrm{p_T}$, electromagnetic isolation, charged track isolation and electron rejection.

5.1.13 Calibration and tagging efficiency

In the following subsections, the effect of mechanical sensor displacement in the tracker is evaluated. Furthermore, methods for measuring the τ-jet energy scale, jet $\rightarrow$ τ mistagging rate, and the τ-jet identification efficiency from data are described.

5.1.13.1 Effects of misalignment on the τ-jet identification

Since many of the τ-jet identification methods are based on charged tracks, it is necessary to understand the effect of sensor displacement in the tracker, i.e. the tracker misalignment, on the performance of these methods. This section evaluates the effects of tracker misalignment on the key τ-jet identification methods which rely on the tracker. The evaluation was performed by comparing the results of the τ-jet identification methods given by the three simulated misalignment scenarios defined in Section 4.4.1. These scenarios are referred to in the following subsections by a number. Scenario 0 corresponds to perfect alignment, Scenario 1 denotes the first data taking scenario (first few 100 pb^{-1} of integrated luminosity), and Scenario 2 stands for the long-term data taking scenario (first few 1 fb^{-1} of integrated luminosity).

The misalignment was simulated with the event samples described in Section 5.1.1. The parameters $\mathrm{R_m}$ = 0.10, $\mathrm{R_s}$ = 0.07, $\mathrm{R_i}$ = 0.40, $\Delta\mathrm{IP_z} < 2$ mm, and $\mathrm{p_T^{min}} > 1$ GeV/c were used for the charged track isolation. The charged track isolation criteria were applied to all the results shown for the misalignment effect evaluation unless otherwise stated.

A standard set of track quality cuts was applied for the tracks in a jet, which passed the charged track isolation criteria, in order to remove badly reconstructed tracks. These track quality cuts are:

- leading track $\mathrm{p_T} \geq$ 10 GeV/c and track $\mathrm{p_T} \geq$ 1 GeV/c for other tracks,
- $\mathrm{n_{hits}} \geq 8$ for leading track and $\mathrm{n_{hits}} \geq 5$ for other tracks,
- $\mathrm{n_{pixel\ hits}} \geq 2$ for all tracks,
- normalized $\chi^2 \leq 10$ for the leading track, and
- normalized hit estimate $\chi^2 \leq 10$ for all hits in all tracks.

The possibility to use tighter cuts for the case with three reconstructed tracks in the signal cone is discussed in more detail in Section 5.1.13.1.2. The benefit of the additional requirements was found to be small and therefore the track quality cuts mentioned above were applied to the cases of both one and three reconstructed tracks in the signal cone of the charged track isolation algorithm. The cut on the hit estimate chi^2 removes those hadronic jets from QCD multi-jet events, in which a single reconstructed track consists of hits coming from different simulated tracks as described in Section 5.1.7. The cut is effective, because this kind of track composition results in abnormally high hit estimate values for the two innermost hits. However, the cut does, not suppress all the hadronic jets of this type. Thus it is necessary to apply an upper limit cut to the transverse impact parameter.

The track quality cuts were applied to all the results of the misalignment effect evaluation, except for the jet and track multiplicities in Sections 5.1.13.1.1 and 5.1.13.1.3. In the following subsections, the effect of misalignment is evaluated first for the jets and tracks and then for the impact parameter and flight path reconstruction.

5.1.13.1.1 Effect of misalignment on the charged track isolation

The algorithm for charged track isolation has been described in Section 5.1.6. Table 5.9 shows an increase in the number of hadronic jets from QCD multi-jet events passing the charged track isolation criteria in the misaligned scenarios compared to the the perfect alignment scenario. The increase in the number of jets passing the charged track isolation criteria was found to be considerably larger in the first data taking scenario (Scenario 1) than in the long-term scenario (Scenario 2). The largest increase was observed to occur at the jet $\mathrm{E_T^{MC}}$ = 50-80 GeV bin.

	Jet $\mathrm{E_T^{MC}}$			
	30-50 GeV	50-80 GeV	80-120 GeV	120-170 GeV
Scenario 1	+61.7±0.6%	+64.5±0.5%	+47.1±0.5%	+27.9±0.5%
Scenario 2	+6.6±0.5%	+6.8±0.5%	+5.6±0.6%	+4.6±0.5%

Table 5.9: The change in the number of hadronic jets from QCD multi-jet events passing charged track isolation criteria in the different misaligned scenarios compared to perfect alignment in different jet $\mathrm{E_T^{MC}}$ bins. The errors shown represent statistical uncertainty. No track quality cuts have been applied.

The increase in the number of hadronic jets passing the charged track isolation criteria is explained by the charged track isolation algorithm topology. When sensor displacement is added, the probability increases for rejecting tracks which are reconstructed in the isolation cone in the perfect alignment scenario. This effect results from the impact parameter z coordinate being more than 2 mm apart from the $\mathrm{IP_z}$ of the leading track. Furthermore, the probability to reconstruct a single track in a hadronic jet from a small number of hits belonging to different simulated tracks increases with sensor displacement, whereas these tracks would not have been reconstructed in the perfect alignment scenario. In these cases, misalignment causes the jet to be interpreted as an isolated jet and thus to pass the charged track isolation criteria. This effect leads to a higher number of jets and consequently events passing the charged track isolation criteria for the scenarios with sensor displacement than for the perfect alignment scenario. The proportion of such hadronic jets with just one track in the isolation cone compared to all hadronic jets is lower for the higher jet $\mathrm{E_T^{MC}}$ bins, because the average number of tracks in a hadronic jet and thus also in the isolation cone increases with jet $\mathrm{E_T^{MC}}$.

The change in the number of genuine τ jets passing the isolation criteria as a result of misalignment was found to be small, because the τ jets contain considerably fewer tracks than hadronic jets, and because the charged track isolation algorithm is designed, a priori, to select the narrow decay signature of the τ jets.

5.1.13.1.2 Background rejection by track quality cuts

Figure 5.33 shows the effect of sensor displacement on the signal and background efficiencies, when track quality cuts are applied after the charged track isolation. The efficiency was calculated by comparing the number of jets surviving the charged track isolation against the number of jets surviving both the charged track isolation and the track quality cuts.

For τ jets with one reconstructed track in the signal cone (top left plot of Fig. 5.33), the failure to pass the track quality cuts is dominated by the $\mathrm{p_T}$ cut on the track. The efficiency for passing the track quality cuts is observed to increase with jet $\mathrm{E_T^{MC}}$, because the average track $\mathrm{p_T}$ increases with jet $\mathrm{E_T^{MC}}$ and hence a larger proportion of jets passes the $\mathrm{p_T}$ cut. When misalignment is introduced, the alignment position error causes the search window for hits to be broader compared to the perfect alignment scenario, which translates into an increase in the efficiency. Some of the increased efficiency can be traced back to 'fake' tracks, in which three simulated charged tracks have been accidentally combined as one reconstructed track. As the jet $\mathrm{E_T^{MC}}$ is increased, the three simulated tracks become more collimated and the probability for this effect is increased.

For hadronic jets with one reconstructed track in the signal cone (top right plot of Fig. 5.33), most of the jets are rejected with the track quality cuts, because their track contains less than eight hits. In the hadronization process of the jets, several relatively long-lived charged particles are produced. Therefore, the observation of just one isolated reconstructed track implies that a number of tracks have been lost in the reconstruction process, but most of their hits exist in the tracker. In such a case, the surviving

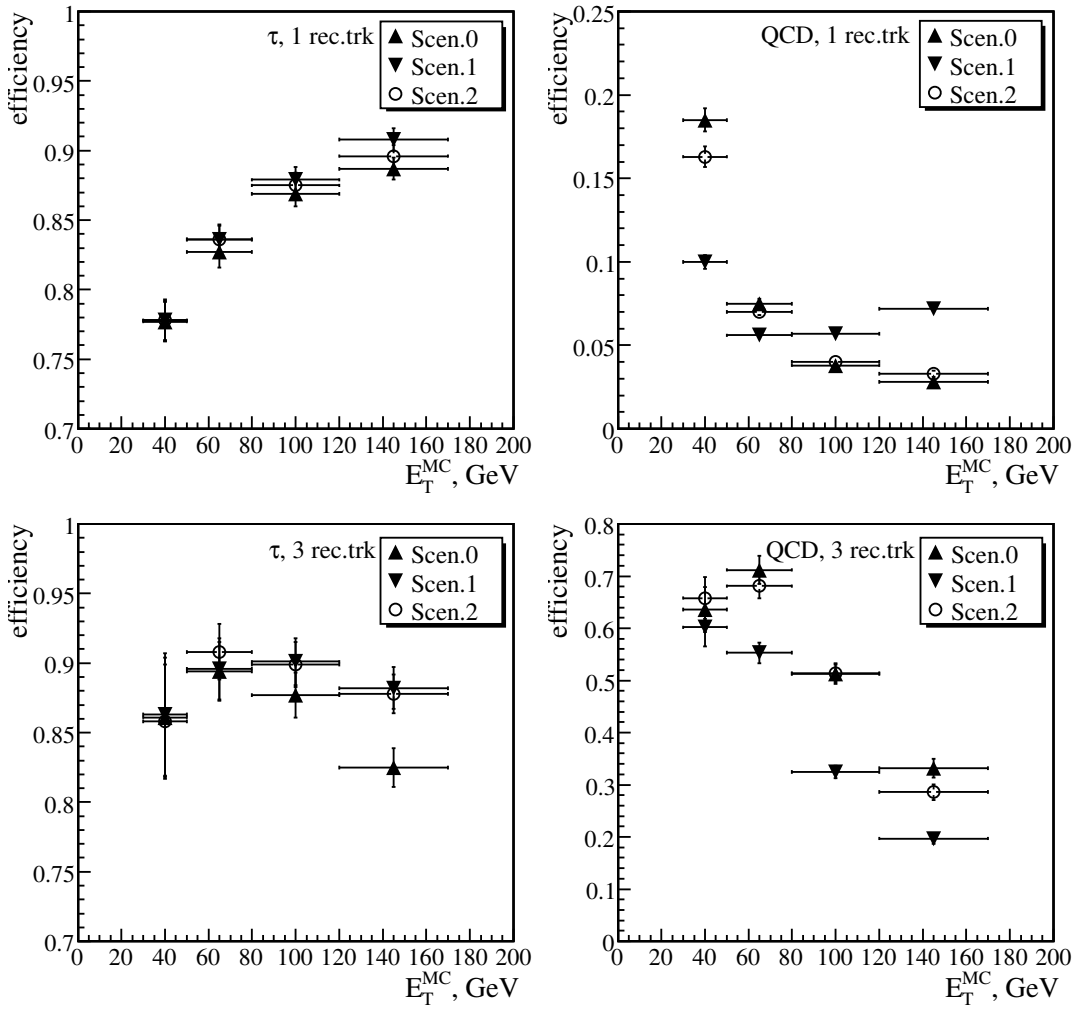

Figure 5.33: The signal and background efficiencies for the different jet $\mathrm{E_T^{MC}}$ bins and misalignment scenarios for passing track quality cuts after charged track isolation. The values are shown for jets with one and three reconstructed tracks in the signal cone. The errors shown represent statistical uncertainty.

track may be built of hits belonging to different simulated tracks (see Section 5.1.13.1.4). Many of these 'fake' tracks consist of a small number of hits and hence most of the jets are rejected as a result of the requirement of having at least eight hits in the track. As the jet $\mathrm{E_T^{MC}}$ is increased, the average number of tracks and consequently the average number of hits populating the tracker increases. Consequently, also the probability for assembling a fake track increases (hit population is more dense) and hence the background rejection increases with the jet $\mathrm{E_T^{MC}}$. When the tracking sensors are displaced, the alignment position error causes the search window for hits in track finding to be broader compared to the perfect alignment scenario. This translates into building more 'fake' tracks than in the perfect alignment scenario. At low jet $\mathrm{E_T^{MC}}$ values, these fake tracks tend to get suppressed since they contain less than eight hits, which is why the rejection of jets by the track quality criteria is increased compared to the perfect alignment scenario when sensor displacement is introduced. As the jet $\mathrm{E_T^{MC}}$ is increased, the number of hits in the fake tracks increases on the average. Hence more jets tend to pass the the track quality cuts for high jet $\mathrm{E_T^{MC}}$ in the misaligned scenarios compared to the perfect alignment scenario.

The efficiency curves for passing the track quality cuts for τ jets with three reconstructed tracks in the signal cone (bottom left plot of Fig. 5.33) behave similarly as for the τ jets with one reconstructed track for low jet $\mathrm{E_T^{MC}}$ values. The cut on the leading track $\mathrm{p_T}$ dominates the inefficiency for the low jet $\mathrm{E_T^{MC}}$ values. Compared to the τ jets with one reconstructed track in the signal cone, the efficiency for passing the track quality cuts is somewhat higher for the τ jets with three reconstructed track in the signal cone, because the probability to have at least one track exceeding the $\mathrm{p_T}$ threshold is increased, if there are more than one reconstructed tracks to choose the leading track from. For high jet $\mathrm{E_T^{MC}}$ values, the efficiency for passing the track quality cuts is observed to decrease as a result of the cut on the hit estimate χ^2. The kinematical cone, which contains the three tracks, shrinks with jet $\mathrm{E_T^{MC}}$ and hence the probability, that the tracks share at least one of the hits, is increased. The shared hits cause the track fit to worsen in the first few hits of the tracks and consequently some of these tracks fail the hit estimate χ^2. This decrease of efficiency is much smaller in the misaligned scenarios than in the perfect alignment scenario, since the adding of the APE in the denominator of the hit estimate χ^2 decreases the χ^2 value. Consequently, because of more relaxed track fitting parameters, less jets are rejected with the track quality cuts for scenarios with sensor displacement compared to the perfect alignment scenario, even though the probability of sharing hits increases with sensor displacement.

For hadronic jets with three reconstructed tracks in the signal cone (Fig. 5.33, lower right), the efficiency for passing the track quality cuts is found to be considerably greater than for hadronic jets with one reconstructed track in the signal cone. This is because the hadronic jets consist of several charged particles and therefore the probability of having reconstructed at least few tracks properly is the greater the more tracks are reconstructed. For low jet $\mathrm{E_T^{MC}}$ values, the cut on the $\mathrm{p_T}$ is the dominating cause for failing the track quality cuts. For high jet $\mathrm{E_T^{MC}}$ values, the probability for the before mentioned 'fake' tracks increases. Therefore, the cut on the hit estimate χ^2 and leading track hit number become the dominant reason for failing the track quality cuts for high jet $\mathrm{E_T^{MC}}$ values. Since the proportion of jets with fake tracks increases with the jet $\mathrm{E_T^{MC}}$, also the rejection increases with jet $\mathrm{E_T^{MC}}$. With sensor displacement, the proportion of jets with fake tracks is increased and therefore the rejection with track quality cuts is increased with sensor displacement.

For the case of three reconstructed tracks in the signal cone, track quality cuts requiring each track to have at least eight hits per track was also investigated. The resulting change in the signal and background efficiencies compared to the standard track quality cuts is shown in Table 5.10. It can be seen from the table, that the tightened track quality requirements suppress at least the same amount or more signal than background events. Therefore, it was decided to demand only the leading track to contain at least eight hits and the second and third tracks were required to have at least five hits as demanded by the track building.

	Signal	QCD bkg.
Scenario 0		
$\mathrm{E_T^{MC}}$ 30-50 GeV	-14±3%	-9±5%
$\mathrm{E_T^{MC}}$ 50-80 GeV	-13±1%	-13±3%
$\mathrm{E_T^{MC}}$ 80-120 GeV	-12±1%	-8±3%
$\mathrm{E_T^{MC}}$ 120-170 GeV	-11±1%	-3±4%
Scenario 1		
$\mathrm{E_T^{MC}}$ 30-50 GeV	-12±3%	-9±4%
$\mathrm{E_T^{MC}}$ 50-80 GeV	-11±1%	-9±3%
$\mathrm{E_T^{MC}}$ 80-120 GeV	-10±1%	-5±3%
$\mathrm{E_T^{MC}}$ 120-170 GeV	-10±1%	-5±3%
Scenario 2		
$\mathrm{E_T^{MC}}$ 30-50 GeV	-13±3%	-11±4%
$\mathrm{E_T^{MC}}$ 50-80 GeV	-12±1%	-11±2%
$\mathrm{E_T^{MC}}$ 80-120 GeV	-11±1%	-9±2%
$\mathrm{E_T^{MC}}$ 120-170 GeV	-12±1%	-3±3%

Table 5.10: The change in the signal and background efficiencies in percent units for different jet $\mathrm{E_T^{MC}}$ bins when requiring all three tracks to contain at least eight hits in addition to the standard track quality cuts. The errors shown represent statistical uncertainty.

5.1.13.1.3 Number of reconstructed tracks

The number of reconstructed tracks in the signal cone is shown in Fig. 5.34 for signal and background events for the different jet $\mathrm{E_T^{MC}}$ bins and misalignment scenarios after applying the charged track isolation and the track quality cuts criteria. The change in the number of reconstructed tracks due to misalignment is small for τ jets which have been MC matched to one- and three-prong final states. For the MC matched three-prong τ jets, the probability for any of the tracks to share some of the first hits in the pixel layers increases both with jet $\mathrm{E_T^{MC}}$ and detector misplacement. This effect was observed to increase the number of jets, where the three simulated tracks are reconstructed as more than three tracks.

For hadronic jets from QCD multi-jet events, little difference is observed between the different misalignment scenarios for the lowest jet $\mathrm{E_T^{MC}}$ bin. As the jet energy is increased, it becomes more likely, that the reconstructed track is built from hits from different simulated tracks. Hence, the number of reconstructed tracks in a jet becomes on average smaller for hadronic jets in the misaligned scenarios compared to the perfect alignment scenario.

5.1.13.1.4 Fake hit rate in tracks

The reconstructed hits of the reconstructed tracks can be matched to simulated hits of simulated tracks to find out which simulated tracks they belong to. This MC matching can be used to define a fake hit rate in the tracks in the following way.

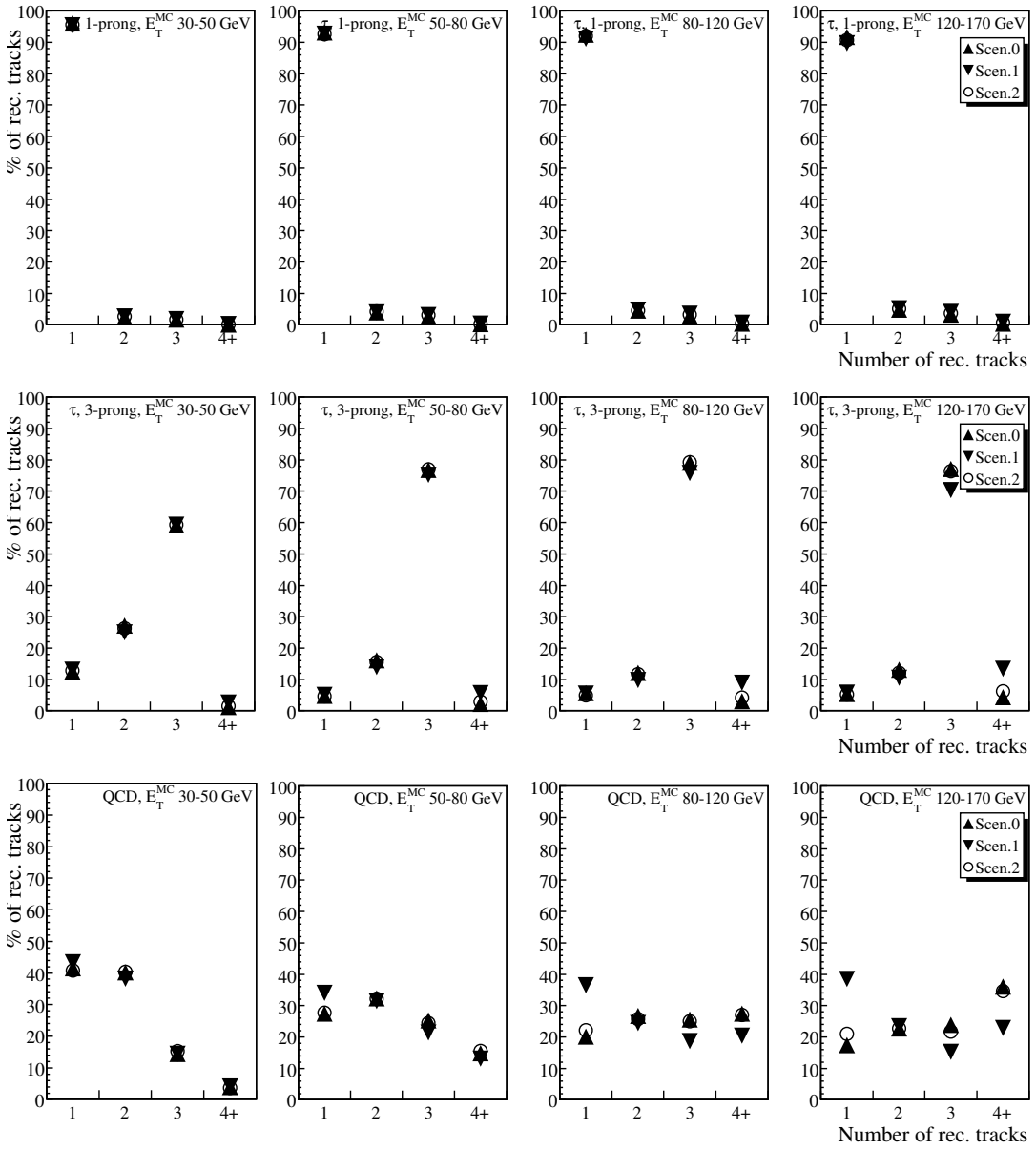

Figure 5.34: The number of reconstructed tracks for MC matched one- (top) and three-prong (middle) τ jets and hadronic jets from QCD multijet events (bottom) for different jet $\mathrm{E_T^{MC}}$ bins and misalignment scenarios after applying the charged track isolation and the track quality cuts. The last bin contains jets with four or more reconstructed tracks.

The simulated track with most matching hits to the reconstructed track was chosen to be the 'true' simulated track which the reconstructed track corresponds to. Each reconstructed hit, which was not associated to a simulated hit belonging to this correct simulated track, was labeled as a 'fake' hit. The fake hit rate of a reconstructed track is then defined as the fraction of these fake hits of all reconstructed hits in the reconstructed track. Thus, a high fake hit rate points out, that the reconstructed track is composed of hits belonging to several simulated tracks. All of the hits are true hits that exist in the tracker, but they might belong to different simulated tracks.

If all reconstructed hits are found to come from different simulated tracks, a fake hit rate of 100% was

assigned to the reconstructed track. For the case of three reconstructed tracks, the fake hit rate was taken as the average of the fake hit rate of the three reconstructed tracks. The simulated tracks contained in addition to the hard event also simulated tracks coming from interaction with tracker material.

The fake hit rate, which was calculated with the definition above, is shown in Figures 5.35 and 5.36 for the different misalignment scenarios for the cases of one and three reconstructed tracks in the signal cone, respectively. The jets were required to pass the charged track isolation and the track quality cuts.

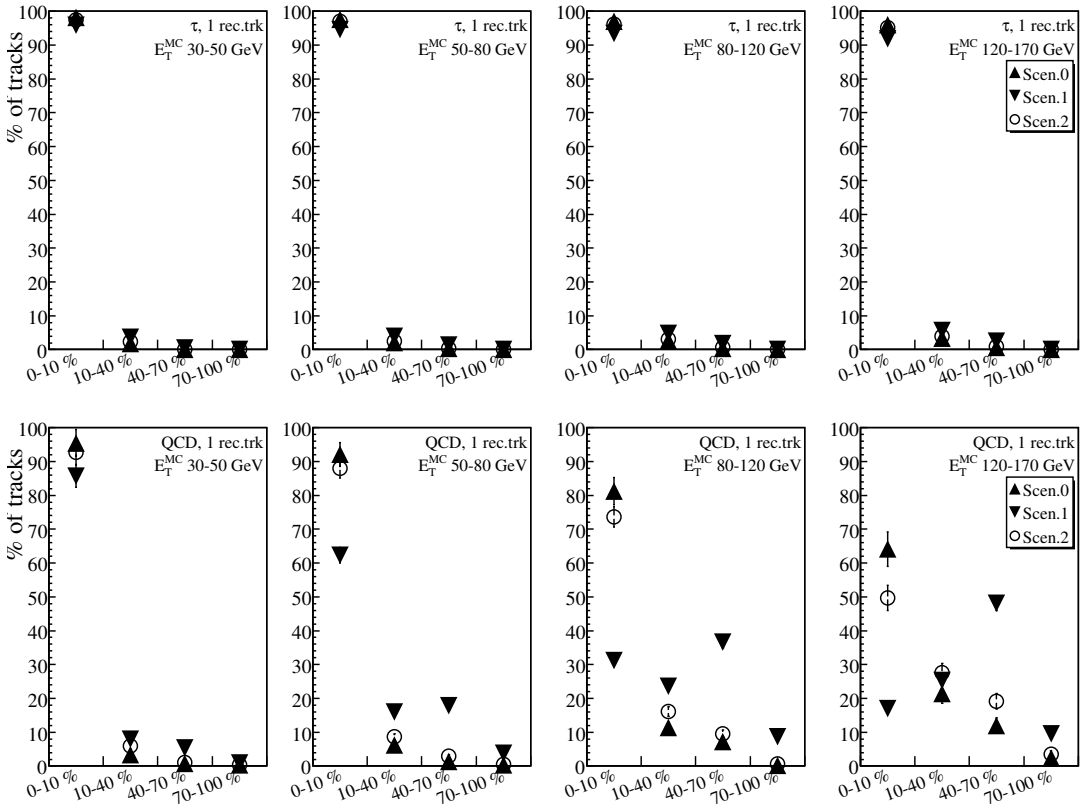

Figure 5.35: The fake hit rate of tracks for different jet $\mathrm{E_T^{MC}}$ bins and misalignment scenarios for τ jets (top) and hadronic jets from QCD multi-jet events (bottom) with one reconstructed track after charged track isolation and the track quality cuts. The errors shown represent statistical uncertainty.

For τ jets with one reconstructed track in the signal cone, the introduction of sensor displacement was found to increase the fake hit rate only slightly, because the tracker was populated with hits coming only from a single track and the hits from tracks produced in occasional interactions with the tracker material. For τ jets with three reconstructed tracks in the signal cone, the fake hit rate was observed to increase slightly more compared to the one reconstructed track case, because the tracks were close to each other and hits could be shared in some pixel detector layers. The overall effect was, however, small. Practically all τ jets were found to have a fake hit rate below 40%.

For hadronic jets from QCD multi-jet events, the fake hit rate was found to increase with both jet $\mathrm{E_T^{MC}}$ and sensor misplacement, since the number of collimated tracks in a hadronic jet is increased with jet $\mathrm{E_T^{MC}}$ and since the enlarging of the sensor position uncertainty enable a broader search cone for the

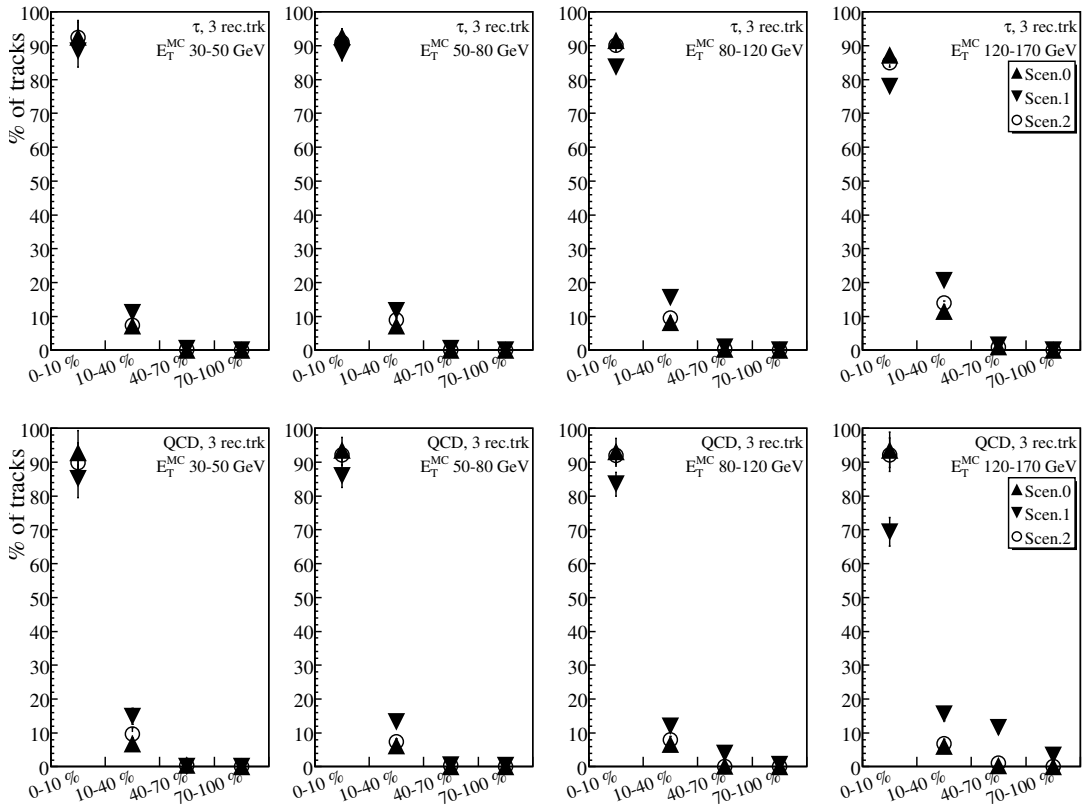

Figure 5.36: The fake hit rate of tracks for different jet $\mathrm{E_T^{MC}}$ bins and misalignment scenarios for τ jets (top) and hadronic jets from QCD multi-jet events (bottom) with three reconstructed track after charged track isolation and the track quality cuts. The errors shown represent statistical uncertainty.

next hit in track finding. In the case of one reconstructed track in the signal cone, the fake hit rate in the tracks was considerably high for the high jet $\mathrm{E_T^{MC}}$ bins. The effect was amplified when sensor displacement was introduced, since the majority of tracks contained several fake hits in the scenarios with sensor displacement. For hadronic jets with three reconstructed tracks in the signal cone, the average fake hit rate was found to be much lower than for the one reconstructed track case. For low jet $\mathrm{E_T^{MC}}$ values there was not much different between the misalignment scenarios, but for high jet $\mathrm{E_T^{MC}}$ values, the fake rate was found to increase in the misaligned scenarios.

The tracks with fake hit rate greater than 20% were observed to be tracks going predominantly in the forward direction ($1.5 < |\eta| < 2.2$) as can be seen in Fig. 5.37. The propagation distance between the pixel sensors and the first silicon strip sensors is long in the forward direction, which increases the possibility that a reconstructed track is composed of hits belonging to different simulated tracks, especially if the search cone for the next hit in track finding is broadened because of the APE.

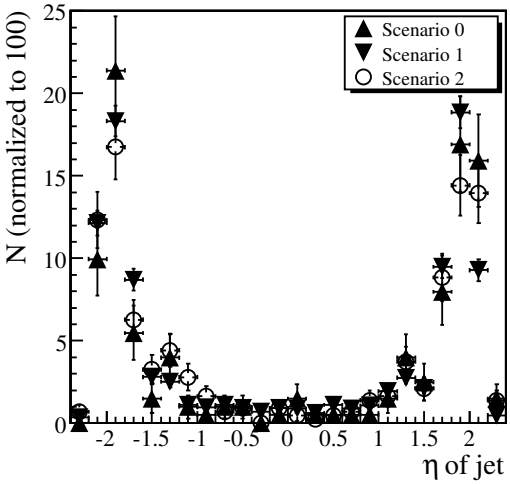

Figure 5.37: The jet η distribution of hadronic jets from QCD multi-jet events with one reconstructed track in the signal cone with at least 20% fake hit rate. The errors shown represent statistical uncertainty.

5.1.13.1.5 Effect of misalignment on the impact parameter

The method for the identification of τ jets with the impact parameter of the leading track has been described in Section 5.1.7. The increase in the number of tracks with high fake hit rate as a result of the misalignment is reflected to the impact parameter measurement. Tracks, which are composed of hits coming from different tracks, have some kinks in them and they are thus expected to have a non-zero impact parameter. This effect becomes especially well visible in the measurement of the impact parameter in the transverse plane (Fig. 5.38). By placing an upper limit cut of $\mathrm{IP_T} < 300$ μm, the tail and hence most of the fake tracks can be efficiently removed independent of the misalignment scenario with minimal signal event losses. Additional cleaning can be achieved with the cut $\mathrm{IP} < 1$ mm in three-dimensional space.

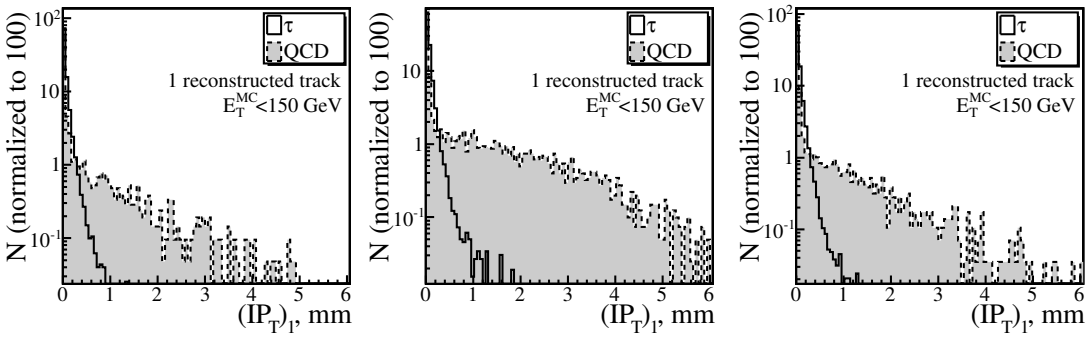

Figure 5.38: The transverse impact parameter distribution for τ jets and hadronic jets from QCD multi-jet events with one reconstructed track in the signal cone in the perfect alignment (left), first data taking (middle), and long-term (right) misalignment scenarios. The histograms have been normalized to 100.

Once the upper limit cuts on the impact parameter were made, the shape of the signal and background

distributions was found to be quite similar (Fig. 5.39) for all the misalignment scenarios. The background distribution was found to be suppressed more than the signal distribution and therefore a lower limit threshold was found to suppress more background than signal. The sensor displacement was found to increase the width of the distributions equally for both the signal and background events. The unsigned three-dimensional impact parameter significance distributions shown in Fig. 5.40 were found to yield little separating power for both signal and background samples independent of sensor displacement. A lower limit cut on the impact parameter was observed to yield background separation in all misalignment scenarios.

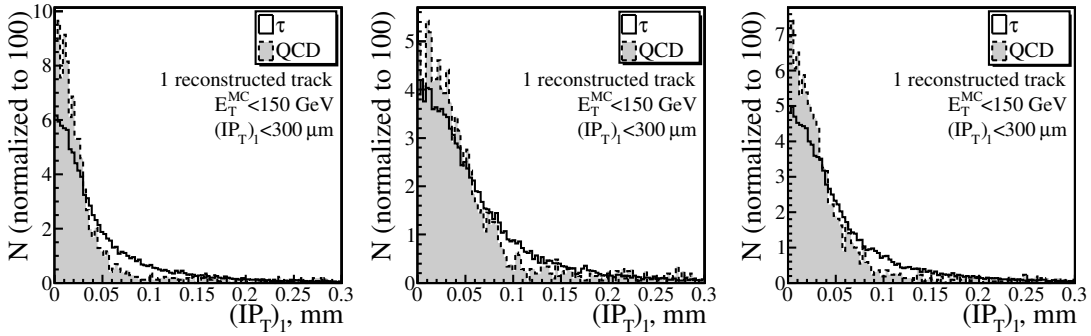

Figure 5.39: The transverse impact parameter distribution for τ jets and hadronic jets from QCD multi-jet events with one reconstructed track in the signal cone in the perfect alignment (left), first data taking (middle), and long-term (right) misalignment scenarios. The plot has been zoomed to $\mathrm{IP_T} < 300$ μm. The histograms have been normalized to 100.

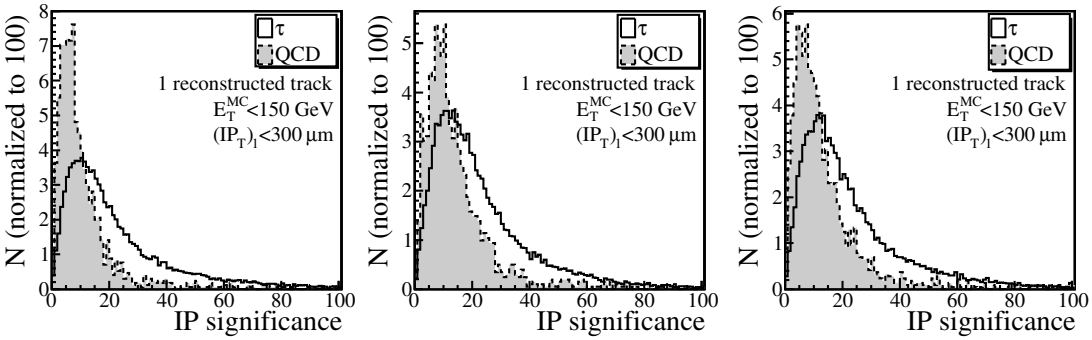

Figure 5.40: The unsigned impact parameter significance distribution for τ jets and hadronic jets from QCD multi-jet events with one reconstructed track in the signal cone in the perfect alignment (left), first data taking (middle), and long-term (right) misalignment scenarios. The histograms have been normalized to 100.

The efficiency curves based on cuts on the impact parameter are shown in Fig. 5.41 for jets with one reconstructed track in the signal cone. The lower limit on the impact parameter significance was varied in the plot between 3 and 14 in steps of 1. The uppermost points correspond to a cut of $\sigma_{\mathrm{IP}} > 3$. The efficiency from removing the 'fake' tracks with the cuts $\mathrm{IP_T} < 300$ μm and $\mathrm{IP} < 1$ mm were included in the efficiency. Because the tail on the transverse impact parameter distributions increases with sensor displacement (Fig. 5.38), and since the tail can be removed, the corresponding background rejection also increases with sensor displacement, as seen in the efficiency curves. The lower limit cut on the

impact parameter significance was found to give separation of signal from background (Fig. 5.41), but the separation increased only slightly with jet $\mathrm{E}_{\mathrm{T}}^{\mathrm{MC}}$.

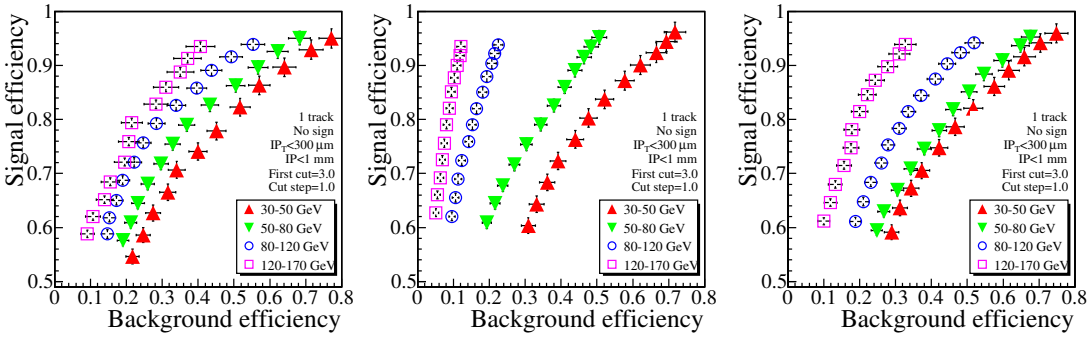

Figure 5.41: The signal and background efficiency curves based on variation of the unsigned full IP significance for one reconstructed track in the signal cone in the different jet $\mathrm{E}_{\mathrm{T}}^{\mathrm{MC}}$ bins. The plots correspond to the different misalignment scenarios (left: perfect alignment, middle: short term scenario, and right: long-term scenario). The cuts $\mathrm{IP}_{\mathrm{T}} < 300$ μm and $\mathrm{IP} < 1$ mm have been included in the efficiency. The error bars indicate the statistical uncertainty.

In the case of three reconstructed tracks in the signal cone, the best background separation was found to be given by the average of the three-dimensional impact parameter significances of the three tracks (Fig. 5.42). The upper limit cut on the transverse and full impact parameters were used only on the leading track, because it was found to have the highest probability of the three tracks of containing hits from other tracks. The separation of the signal and background distributions with the lower limit cut was found to be better in the case of three reconstructed tracks in the signal cone compared to the one track case. On the other hand, the amount of tracks composed from hits belonging to different simulated tracks was considerably smaller in the three track case compared to the one track case. Consequently, the background rejection resulting from fake hit cleaning with the upper limit cut on the impact parameter was found to be smaller in the three track case than in the one track case.

Having three reconstructed tracks in the signal cone was found to provide a more accurate determination of the sign of the impact parameter, i.e. whether the impact parameter vector points to the same direction as the jet, than in the single track case. The positive sign requirement was observed to reduce the signal efficiency by about 10-15% while rejecting 35-45% of the background.

The signal and background efficiency curves based on a lower limit cut on three-dimensional signed impact parameter significance are shown in Fig. 5.43 for the different misalignment scenarios. The curves were calculated by varying the signed impact parameter significance from 3 to 13 in steps of 1, where the uppermost points correspond to the $\sigma_{\mathrm{IP}} > 3$ cut. The efficiency of the upper limit cut on the transverse and three-dimensional impact parameters on the leading track has been included in the curves. The absence of a large number of tracks with a high fake hit rate was found to make the background rejection less dependent on jet $\mathrm{E}_{\mathrm{T}}^{\mathrm{MC}}$ and to cause the background rejection to remain almost independent of sensor displacement.

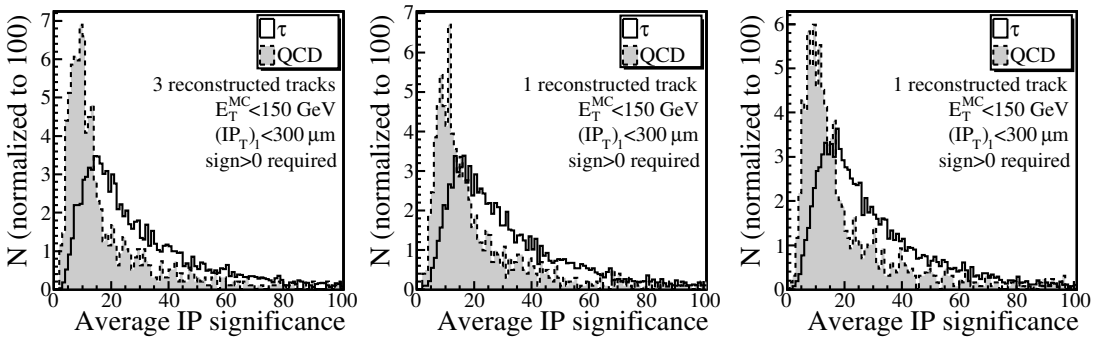

Figure 5.42: The signed average IP significance distribution for τ jets and hadronic jets from QCD multi-jet events with three reconstructed tracks in the signal cone in the perfect alignment (left), first data taking (middle), and long term (right) misalignment scenarios. The histograms have been normalized to 100.

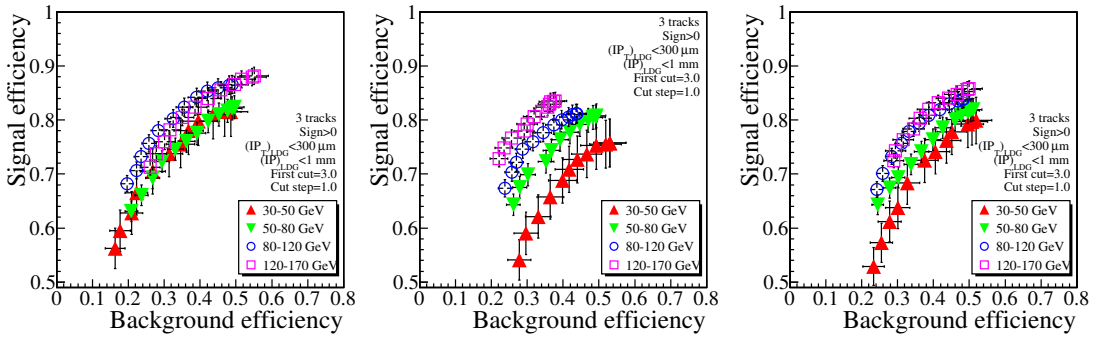

Figure 5.43: The efficiency curves for the signed average of full IP significance for three reconstructed tracks in the signal cone in the different jet $\mathrm{E_T^{MC}}$ bins. The plots correspond to the different misalignment scenarios (left: perfect alignment, middle: short term scenario, and right: long-term scenario). The cuts $\mathrm{IP_T} < 300$ μm and $\mathrm{IP} < 1$ mm were applied to the leading track only and are included in the efficiency.

5.1.13.1.6 Effect of misalignment on the τ sec. vertex reconstruction

The reconstruction of the secondary vertex for τ jets is described in Section 5.1.8.1. To fit the secondary vertex (SV), i.e. the τ decay vertex, three isolated tracks passing the track quality cuts were required. The fit of the secondary vertex was done with the Kalman vertex fitter (KVF) described in Section 4.3.3.2, because it is the most robust of the vertex fitters currently available. Also the adaptive vertex fitter (AVF) and the trimmed Kalman fitter (TKF) were considered and found to give similar results. To keep the studied sample clean, the simulated transverse energy of the jets was required to be below 150 GeV and the reconstructed secondary vertex was required to be within the radius of the first pixel sensor (i.e. 40 mm) layer, unless otherwise stated.

The residual and pull distributions of the reconstructed secondary vertices were used to evaluate the secondary vertex reconstruction performance. The definition of the residual and pull quantities as well as coordinate frame definitions are described in Section 5.1.8.1. The sum of two Gaussian curves was fitted to the residual distributions to estimate the resolution of the secondary vertex fit from the σ of the center Gaussian. Similarly, a double Gaussian curve was also fitted to the pull distributions.

The residual distributions in the global x and z coordinates were observed to contain large tails (Fig. 5.44). The tails were found to come from jets in which at least two of the three tracks are very collimated and additionally from forward jets for the z coordinate. The resolution of the global x coordinate was found to worsen, when sensor displacement was introduced. The resolution of the z coordinate was found to differ from the resolution in the x coordinate, since the tails were larger in the global z direction influencing hence the fit. The bias of the fitted secondary vertex resolution was observed to be statistically small. The pull values were found to decrease with sensor displacement in both x and z directions (Fig. 5.45), which can be interpreted to be caused by an overestimation of the APE in the misaligned scenarios. As mentioned in Section 4.4.1 the APE was added quadratically to the measurement uncertainty.

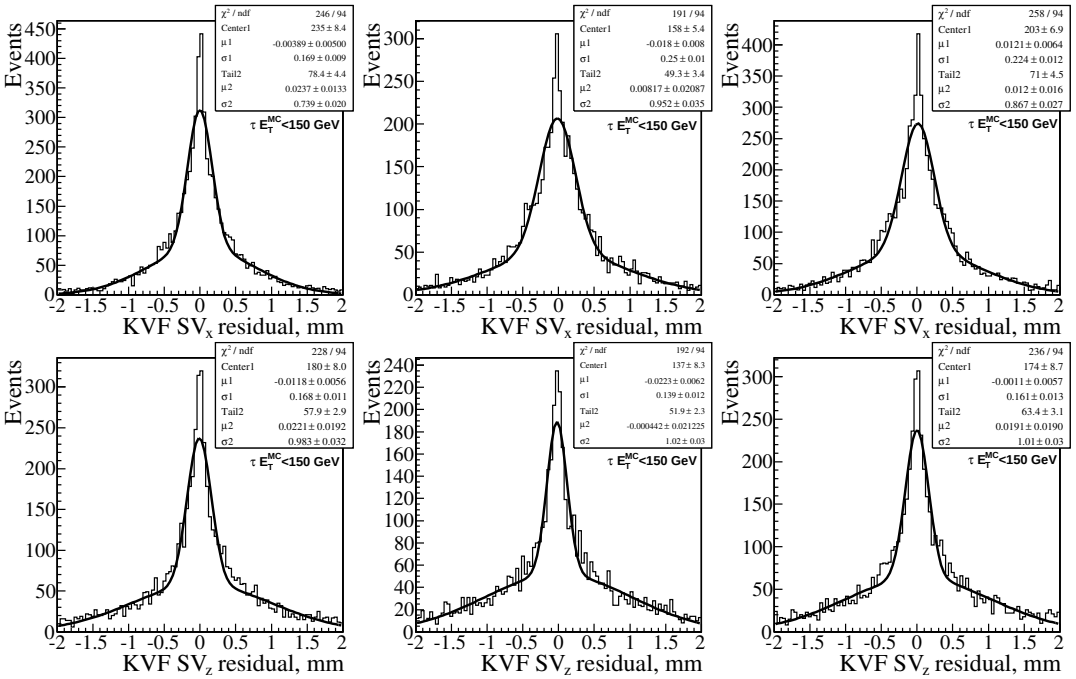

Figure 5.44: The residual of the secondary vertex in the x coordinate (upper plots) and in the z coordinate (lower plots) fitted with the Kalman vertex fitter. The left, middle, and right-hand plots correspond to the perfect alignment, first data taking, and long-term misalignment scenarios, respectively.

The residual distributions in the components transverse to (i) and along (k) the jet axis were found to have smaller tails compared to the residuals in the global coordinate system (Fig. 5.46). Similarly as for the global coordinates, the tails were established to consist of jets where at least two of the three tracks are very collimated. The asymmetry of the tails was found to be caused by tracks having shared hits in the innermost pixel sensor layers. The resolution of the secondary vertex fit was observed to become worse in both i and k directions with sensor displacement, as expected. The pull in the k coordinate was observed to decrease with sensor displacement (Fig. 5.47) as a result of an overestimation of the APE in the misaligned scenarios. In the i direction, the overestimation of the APE was found to be not enough to compensate for the worsening of the residual with sensor displacement and hence the pull distribution are observed to become broader with sensor displacement in the i direction.

The key parameters of the residual and pull distributions are summarized in Table 5.11 for the global x

and z directions and in Table 5.12 in the directions transverse to and along the jet axis (i and k directions). The resolution of the fitted vertex was observed to remain more or less a constant with respect to jet $\mathrm{E_T^{MC}}$ in the direction transverse to the jet axis (left plot of Fig. 5.48). The resolution was found to worsen from $\sim$10 μm of the perfect alignment scenario to $\sim$40 μm in the long-term scenario. In the direction along the jet axis, the resolution of the fitted vertex was found to have a strong dependence on the jet $\mathrm{E_T^{MC}}$ as can be seen in the right plot of Fig. 5.48. The resolution was observed to worsen considerably as a function of sensor displacement.

5.1.13.1.7 Effect of misalignment on the flight path reconstruction

Once the secondary vertex is fitted and since the primary vertex is known, the τ decay length, i.e. the flight path, may be reconstructed. If the flight path vector is pointing opposite to the jet direction, the flight path is assigned a minus sign. Because the three isolated tracks coming from the τ decay are very collimated as a result of the boost effect and because the τ travels some distance before decaying, there is a good chance, that not all of the tracks are separated at the innermost pixel layer. Because of the limited granularity of the pixel sensors, the track may hence share one or more hits. The probability for shared hits is increased with jet energy, because both the decay length and the collimation are increased with jet energy.

If at least one hit was shared between at least two of the three tracks, the secondary vertex position was

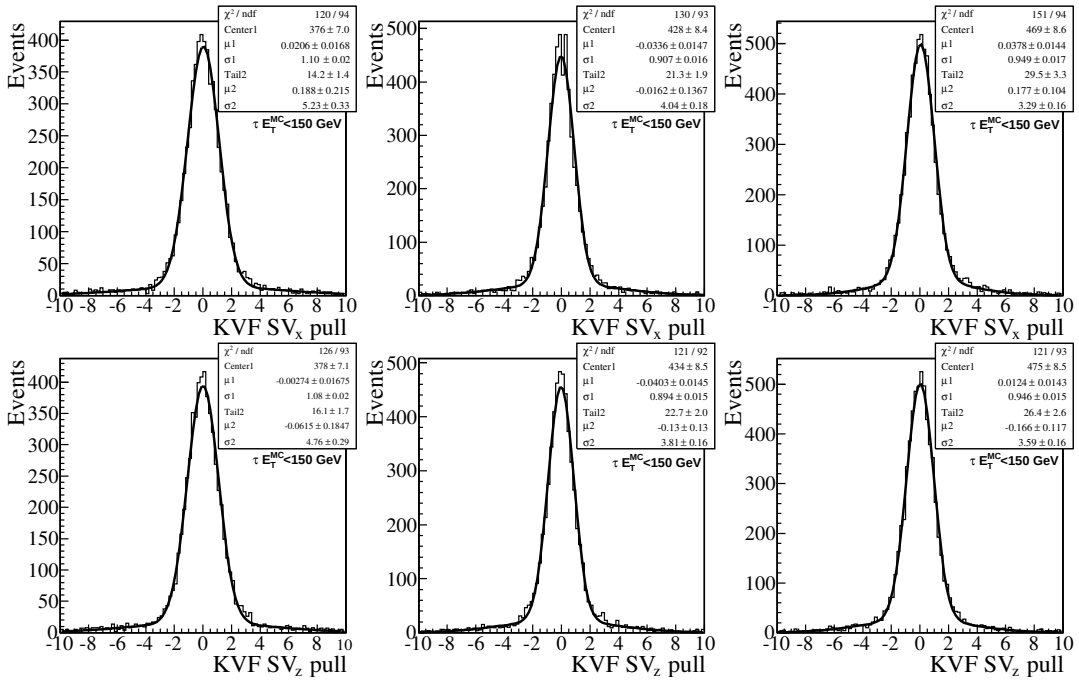

Figure 5.45: The pull of the secondary vertex in the x coordinate (upper plots) and in the z coordinate (lower plots) fitted with the Kalman vertex fitter. The left, middle, and right-hand plots correspond to the perfect alignment, first data taking, and long-term misalignment scenarios, respectively.

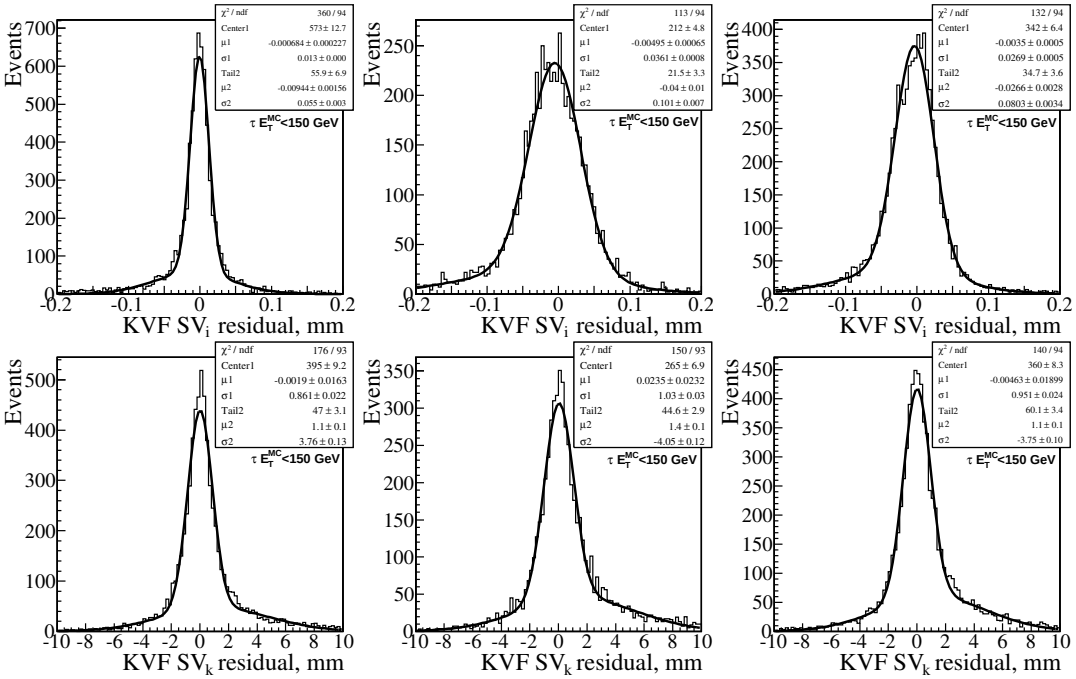

Figure 5.46: The residual of the secondary vertex in the i coordinate (upper plots) and in the k coordinate (lower plots) fitted with the Kalman vertex fitter. The left, middle, and right-hand plots correspond to the perfect alignment, first data taking, and long-term misalignment scenarios, respectively.

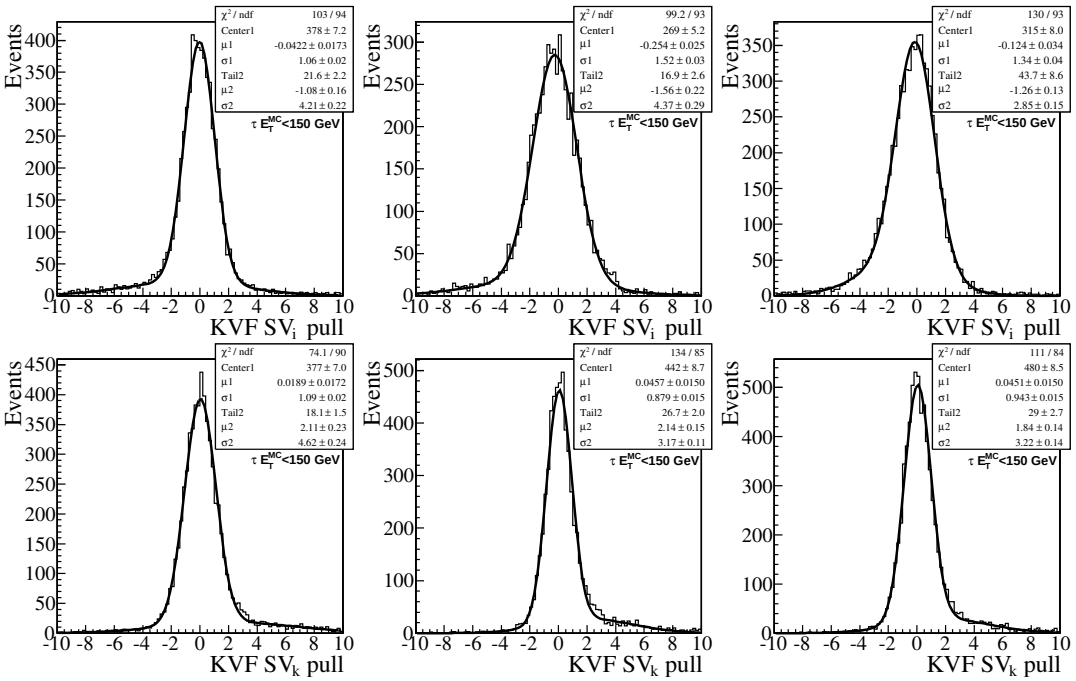

Figure 5.47: The pull of the secondary vertex transverse to the τ-jet axis (i coordinate, upper plots) and along the τ-jet axis (k coordinate, lower plots) fitted with the Kalman vertex fitter. The left, middle, and right-hand plots correspond to the perfect alignment, first data taking, and long-term misalignment scenarios, respectively.

	Scenario 0	Scenario 1	Scenario 2
		x **coordinate**	
χ^2/ndf	2.62	2.03	2.75
σ, μm	169$\pm$9	250$\pm$12	224$\pm$12
Bias, mm	-4$\pm$5	-18$\pm$8	12$\pm$6
Inside $\pm2\sigma$, %	51.6	54.9	54.4
Pull	1.10$\pm$0.02	0.91$\pm$0.02	0.95$\pm$0.02
		z **coordinate**	
χ^2/ndf	2.43	2.04	2.51
σ, μm	168$\pm$11	139$\pm$12	161$\pm$13
Bias, mm	-12$\pm$6	-22$\pm$6	-1$\pm$6
Inside $\pm2\sigma$, %	38.8	28.4	33.8
Pull	1.08$\pm$0.02	0.89$\pm$0.02	0.95$\pm$0.02

Table 5.11: The key parameters of the secondary vertex fit in x and z direction for the τ sample in the perfect alignment, first data taking, and long-term misalignment scenarios.

	Scenario 0	Scenario 1	Scenario 2
		i **coordinate**	
χ^2/ndf	3.83	1.20	1.41
σ, μm	13.0$\pm$0.4	36.1$\pm$0.8	26.9$\pm$0.5
Bias, μm	-0.7$\pm$0.2	-5.0$\pm$0.7	-3.5$\pm$0.5
Inside $\pm2\sigma$, %	71.0	80.6	80.2
Pull	1.06$\pm$0.02	1.52$\pm$0.03	1.34$\pm$0.04
		k **coordinate**	
χ^2/ndf	1.90	1.62	1.49
σ, μm	861$\pm$22	1033$\pm$29	951$\pm$24
Bias, μm	-2$\pm$16	24$\pm$23	-5$\pm$19
Inside $\pm2\sigma$, %	66.7	62.4	64.5
Pull	1.09$\pm$0.02	0.88$\pm$0.01	0.94$\pm$0.01

Table 5.12: The key parameters of the secondary vertex fit in i and k direction for the τ sample in the perfect alignment, first data taking, and long-term misalignment scenarios.

found to be biased towards the first pixel sensor layer in the transverse plane of the global coordinate system. Furthermore, if at least three hits were shared between at least two of the three tracks, the fit of the secondary vertex was observed to be likely to converge to the position of the first pixel sensor layer. The latter effect is seen in the distribution of the reconstructed transverse flight path length as narrow peaks compatible with the pixel sensor radii (Fig. 5.49). When the sensors are displaced, the position uncertainty of the sensors is increased and the narrow peaks are observed to become smeared. The 'resolution' of these peaks is dominated by the precision of the pixel sensors. Since the APE is more overestimated in the first data scenario compared to the long-term scenario, the weight of the pixel hits compared in track fitting is larger in the first data scenario than in the long term scenario. Therefore, the

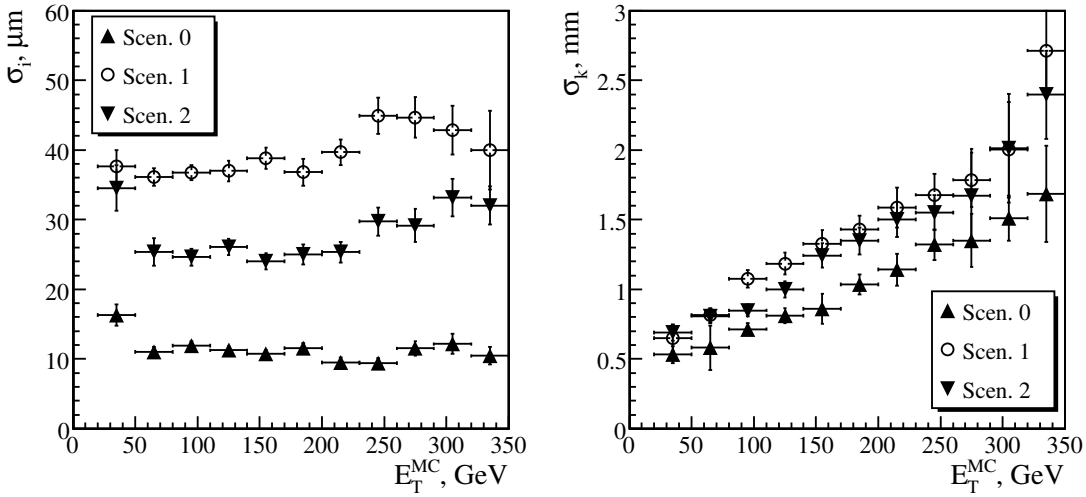

Figure 5.48: The resolution of the secondary vertex fit measured from for the component transverse to jet axis (left plot) and for the component along the jet axis (right plot). Kalman vertex fitter was used for the fitting of the vertex.

peaks were found to be narrower in the first data scenario than in the long-term scenario. The number of jets with tracks sharing hits was found to increase in the scenarios with sensor displacement compared to the perfect alignment scenario, because the hard assignment of hits in the absence of position uncertainty proved to be a too tight constraint for track fitting and some tracks were therefore lost.

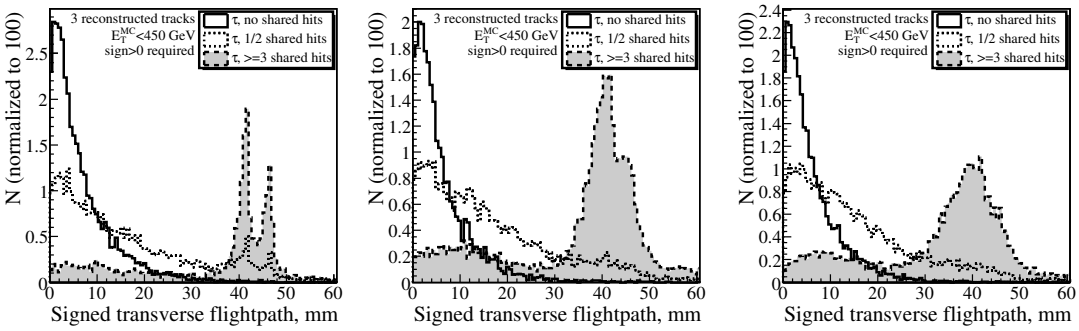

Figure 5.49: The signed transverse flight path length distribution for τ jets with no (solid), one or two (dotted), and at least three (dashed) shared hits in the perfect alignment (left), first data taking (middle), and long-term (right) misalignment scenarios. In each scenario, the histograms have been normalized to the total number of events in the scenario independent of the number of shared hits and the maximum is scaled to 100. Kalman vertex fitter was used to fit the secondary vertex.

The signed three-dimensional flight path significance calculated with the definition in Section 5.1.8 is shown in Fig. 5.50. The full covariance matrices of both the primary and secondary vertices were used in the significance calculation. It was observed, that increasing the sensor displacement affected the shape of the signal and background distributions only minimally. To minimize the effect of shared hits, only jets with reconstructed transverse flight path below 35 mm were considered. The number of jets with $\mathrm{E}_{\mathrm{T}}^{\mathrm{MC}} < 170$ GeV and reconstructed transverse flight path length above 35 mm was about 2.5-5% and 1.3-5% depending on the misalignment scenario for τ jets and hadronic jets from QCD multi-jet events,

respectively.

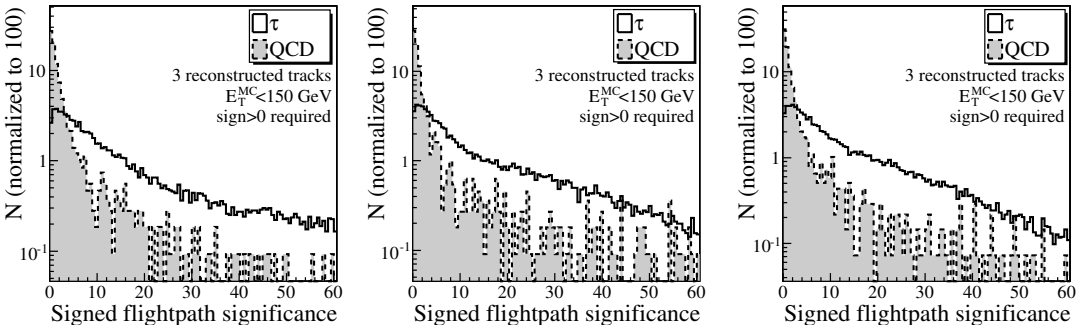

Figure 5.50: The signed full flight path length significance distribution for τ jets and hadronic jets from QCD multi-jet events in the perfect alignment (left), first data taking (middle), and long-term (right) misalignment scenarios. The histograms have been normalized to 100.

Since the signed full flight path significance distributions were found to depend only minimally on the sensor displacement, the signal and background efficiency curves shown in Fig. 5.51 for cutting on this variable were found to differ only little between the misalignment scenarios. The efficiency curves were calculated by varying the cut on the signed 3D flight path significance from 0 to 5 in steps of 0.5. In all misalignment scenarios, the charged track isolation criteria was observed to suppress efficiently the b- and c-jets, which could fake a τ decay because of a non-zero flight path length.

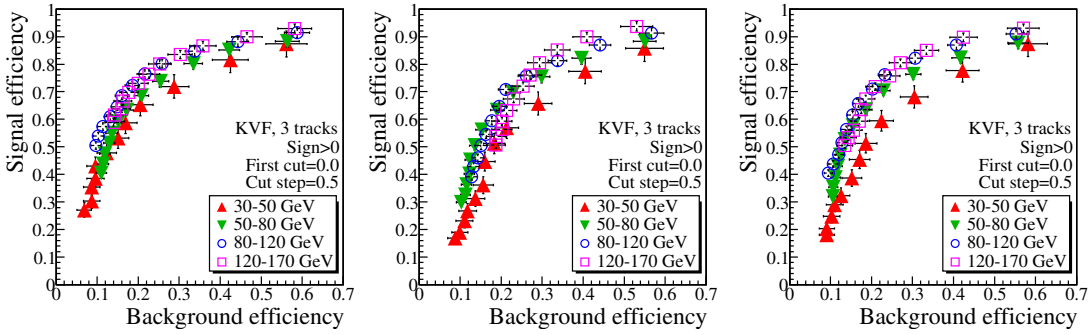

Figure 5.51: The signal and background efficiency curves for the signed full flight path length significance for the different jet $\mathrm{E_T^{MC}}$ bins with three reconstructed tracks in the signal cone in the perfect alignment (left), first data taking (middle), and long-term (right) misalignment scenarios.

5.1.13.1.8 Summary of the effects of misalignment

The introduction of sensor displacement was found to increase the number of jets passing the charged track isolation criteria in the QCD multi-jet events. This increase was found to be mostly caused by 'fake' tracks, which are composed of hits belonging to different simulated tracks. The probability of the occurrence of such 'fake' tracks was found to increase with both jet $\mathrm{E_T^{MC}}$ and sensor displacement. Placing an upper limit cut on the transverse and/or three-dimensional impact parameter of the tracks was

found to reject these 'fake' tracks efficiently and hence to counteract on the increase in the background resulting from the 'fake' tracks.

Hence, the overall effects of misalignment on track-based τ-jet identification methods can be approximated from the efficiency plots for the τ flight path reconstruction as being $\sim$5-10% for the first data taking scenario and slightly less for the long-term scenario. To determine more exact estimates for systematic uncertainty treatment, the analysis would have to be repeated with careful cleaning of the 'fake' tracks before assessing the effect of misalignment. After the publication of the physics TDRs and after the transition from the ORCA reconstruction and analysis software to CMSSW, the occurrence of the 'fake' tracks encountered in this analysis has been greatly reduced.

5.1.13.2 Tau-jet energy scale and calibration with calorimeter

The γ + jet events, where the jet passes the τ-jet identification criteria and thus becomes a τ-like jet, can be used to setup the initial τ-jet energy scale from the real data. Figure 5.52 shows the mean value of the distribution of the ratio $\mathrm{E_T}^{\mathrm{reco}}/\mathrm{E_T^{MC}}$ for the unpreselected hadronic jets from QCD multi-jet events, τ-like hadronic jets from QCD multi-jet events and the real τ jets. Both the hadronic and the τ jets were reconstructed with the iterative cone algorithm described in Section 4.3.1.1 with a cone size of $\mathrm{R_{jet}} = 0.4$. The same cone size was used to evaluate the true transverse energy $\mathrm{E_T^{MC}}$ of the simulated hadronic jets from QCD multi-jet events. The τ-jet identification included the electromagnetic and the charged track isolation with the parameters $\mathrm{P_{isol}^{cut}} < 5$ GeV, $\mathrm{R_i} = 0.4$, and $\mathrm{R_s} = 0.07$. One or three tracks were required to be in the signal cone and the transverse momentum of the leading track was required to exceed 10 GeV/c. It can be seen in Fig. 5.52, that the τ-like hadronic jets from QCD multi-jet events produce a higher calorimeter response than the unpreselected hadronic jets from QCD multi-jet events. The calorimeter response of the τ-like hadronic jets was found to be only 5-10% smaller than the response of the real τ jets. More studies are needed to understand the sources of the remaining difference and the calibration uncertainties.

Another method to evaluate the τ-jet energy scale with the data is to use the $\mathrm{Z} \rightarrow \tau\tau \rightarrow \ell$+jet events and to reconstruct the Z mass peak. This method, however has two disadvantages: the background contamination and the uncertainty of the $\mathrm{E_T^{miss}}$ measurement.

5.1.13.3 Measurement of the jet $\rightarrow \tau$ misidentification from the data

The measurement of the jet $\rightarrow \tau$ misidentification rate can be done with the γ + jet events used for the calibrating the calorimeter response. The advantage of these events is, that the jet energy is known with high precision. About 10^5 of such events are expected for integrated luminosity of 10 fb^{-1} for each $\mathrm{E_T^{\gamma}}$ bin of the size of $0.1 \times \mathrm{E_T^{\gamma}}$, with $|\eta^{\mathrm{jet}}| < 3$ and $\mathrm{E_T^{\gamma}}$ in the interval between 30 and 300 GeV. The mistagging rate can be evaluated then as a fraction of events where the jet has passed the τ-jet identification criteria. Taking into account the jet rejection factor, for example, with the charged track isolation and the τ mass

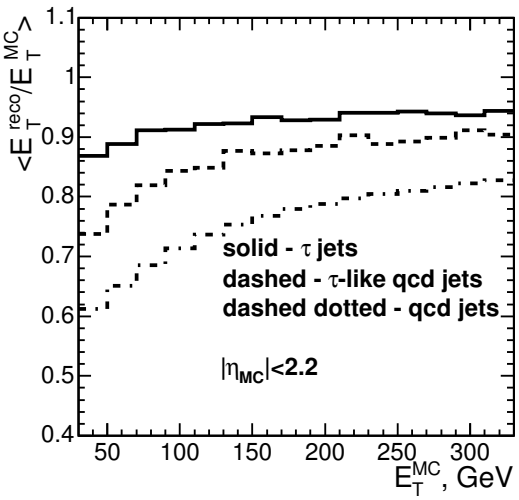

Figure 5.52: The mean value of the ratio $\mathrm{E_T^{reco}}/\mathrm{E_T^{MC}}$ for the hadronic jets from QCD multi-jet events without preselection (dashed-dotted line), τ-like hadronic jets from QCD multi-jet events (dashed line), and the real τ jets (solid line) as a function of $\mathrm{E_T^{MC}}$.

reconstruction (evaluated from the right plot of Fig. 5.12 for $\mathrm{R_i} = 0.4$ and $\mathrm{R_s} = 0.04$, and for the results of Table 5.7) one could expect a 4-10% uncertainty in the estimated mistagging rate per energy bin in the jet $\mathrm{E_T}$ interval of 30-150 GeV with integrated luminosity of 10 fb^{-1}.

The difficulty of the γ + jet method is, that relatively few of the jets have $\mathrm{E_T}$ in excess of 100 GeV needed for example for the $\mathrm{H, A} \rightarrow \tau\tau$ and $\mathrm{H}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ searches. After the publication of the physics TDRs, it has been suggested, that the misidentification rate of jet $\rightarrow \tau$ could also be estimated with clean di-jet or quad-jet events from prescaled jet triggers. In such approach, τ identification is applied on a hadronic jet randomly selected from the QCD di/quad-jet event.

The jet $\rightarrow \tau$ misidentification rate is known to depend on the jet flavor and on the preselections needed for the analysis. Since the γ + jet process produces quark jets, a correction is needed to obtain the contents of hadronic jets from QCD multi-jet events, which are mainly gluon jets. Such correction is not needed for the di/quad-jet events.

5.1.13.4 Measurement of the τ identification efficiency from the data

The τ-jet identification efficiency can be evaluated and compared with simulations by evaluating the ratio of $\mathrm{Z} \rightarrow \tau\tau \rightarrow \mu+$ jet and $\mathrm{Z} \rightarrow \mu\mu$ events selected with the single muon trigger. The reconstruction efficiency of the second muon in the $\mathrm{Z} \rightarrow \mu\mu$ events is assumed to be known. The preliminary estimates were obtained based on the search for MSSM $\mathrm{H, A} \rightarrow \tau\tau \rightarrow \mu+$ jet decay channel described in Ref. [112] by applying the same event selections as for the signal selection of $\mathrm{H, A} \rightarrow \tau\tau \rightarrow$ jets decay channel to be discussed in Chapter 6, but without the b tagging and central jet veto. The systematic uncertainty in the selection cuts on $\mathrm{m_T}(\ell, \mathrm{E_T^{miss}})$, $\mathrm{E_T^{\tau\ jet}}$, and $\mathrm{E_T}$ of the reconstructed neutrinos were

taken into account, as well as the uncertainty of the background evaluation.

With an integrated luminosity of 30 fb^{-1} the total uncertainty of the τ identification efficiency is expected to be between 4 and 5%.

5.1.14 Further developments after the publication of the physics TDRs

Since the publication of the physics TDRs, the main focus of development has been in refining the trigger for the selection of $\mathrm{Z} \rightarrow \tau\tau$ and $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decays at LHC startup luminosities. Since the τ jets are relatively soft in these channels, the trigger thresholds have been lowered and the matching cone size in the charged track isolation algorithm has been increased. The single and double τ trigger strategies have, however, remained the same as what has been described above.

For the offline τ-jet identification, the biggest improvement after the publication of the physics TDRs comes from a holistic approach to reconstruct the τ jets on the level of individual particles with the particle flow method. Although the main τ identification methods described earlier in this chapter remain the same in the particle flow method, it improves the accuracy of both the τ jet energy resolution and the $\mathrm{E}_{\mathrm{T}}^{\mathrm{miss}}$ measurement.

Minor improvements to the τ-jet identification methods include the possibility to use a signal cone, which shrinks depending on the jet E_{T} for the charged track isolation algorithm. Furthermore, the electromagnetic isolation algorithm has been rigged to increase the maximum allowed electromagnetic energy as a function of the jet E_{T}. Furthermore, efforts have been made to combine the separate τ identification methods into a single likelihood or neural net based discriminator. In the following, the particle flow method is described in more detail.

5.1.14.1 The particle flow method

The particle flow method was developed to reconstruct τ-jets accurately and efficiently. The goal of the particle flow algorithm [113] is to provide a complete event description at the level of individually reconstructed particles, including non-isolated and particles, which are merged as a result of the detector granularity. The complete list of particles is then in turn used to derive composite physics objects such as jets or missing E_{T}, which improves the performance of these physics objects compared to the performance obtained with previous algorithms. The identification of different particles coming from the τ decay is done by progressively associating track, ECAL, and HCAL clusters according to the signatures of different particle types. The associations are used to restrict the number of different possible combinations for the so far unreconstructed particles. Multivariate classifiers are used to make the association in some particle types. Finally, the reconstructed particles are clustered into a jet with a cone algorithm [114] in an η,ϕ cone of radius 0.5.

The particle flow algorithm uses the iterative tracking method described in Section 4.3.2.3. The iterative

tracking approach enables the particle flow algorithm also to reconstruct secondary processes occurring in the event. After the first set of track seeds have been removed from the seed list, the patterns of remaining seeds are examined and checked against signatures of nuclear interactions, photon conversions or the decay of unstable hadrons. For example, if the hit density increases along a reconstructed track, a nuclear interaction is interpreted to have occurred.

5.2 Identification of b jets

The identification of b jets [52] is important in many interesting physics channels which contain Higgs bosons, top quarks, or supersymmetric particles. In the analysis of the $\mathrm{H, A} \to \tau\tau$ decay channel, the dominant production mode $\mathrm{gg} \to \mathrm{b\bar{b}H(A)}$ features the production of two associated b jets, whereas in the analysis of the $\mathrm{H}^{\pm} \to \tau^{\pm}\nu_{\tau}$ decay channel, the production mode $\mathrm{gg} \to \mathrm{tbH}^{\pm}$ features the production of one associated b jet with a second b jet coming from the associated top quark decay. Therefore, the identification of at least one of these b jets is a key part of the analysis of the $\mathrm{H, A} \to \tau\tau$ and $\mathrm{H}^{\pm} \to \tau^{\pm}\nu_{\tau}$ channels.

The b jets are different from lighter quark flavor or gluon jets, because they contain b hadrons. The b hadrons are relatively heavy ($\mathrm{m_B} \sim 5.8$ GeV/c^2) and they have a relatively long life-time ($c\tau \sim 420$ μm), which together with the boost effect causes them to have a reconstructable secondary vertex and a non-zero impact parameter. The decay signature of the b hadrons may also involve a reconstructable tertiary vertex, since the b hadrons tend to decay to c hadrons, which can have a long enough life-time to be reconstructable. Since the b hadrons produce a decay chain to lighter quark flavors, the b jets tend to contain a high charged track multiplicity. The decay to lighter flavors occurs in steps through virtual W bosons, which causes $\sim$20% of the b jets to have muons or electrons in the final state coming from the leptonic W decay(s). The background processes, which might mimic a b jet are the lighter flavored jets as well as gluon jets. The most difficult background is constituted by the c jets, since they can produce secondary vertices.

The most robust algorithms used at CMS for the identification, i.e. tagging, of the b jets from other hadronic jets are described in Sections 5.2.1-5.2.4. The strategy for calibrating the b tagging from data is discussed in Section 5.2.5

5.2.1 Track counting algorithm

The track counting b-tagging algorithm [115] exploits the three-dimensional impact parameters of tracks in b jets. The three-dimensional impact parameter of a track is given a minus sign, if it is reconstructed behind the primary vertex with respect to the jet direction. The algorithm starts by constructing a list of the reconstructed tracks with distance of closest approach to the jet axis below 0.7 mm. Then, this track collection is sorted in decreasing order by three-dimensional impact parameter significance, i.e. the three-

dimensional impact parameter value divided by its measurement uncertainty. The discriminator value of the algorithm is defined as the three-dimensional impact parameter significance of the N^{th} track. Jets, in which the discriminator value exceeds a chosen cut value, are declared as b jets.

5.2.2 Probability based algorithm

Also the probability based b-tagging algorithm [115] is based on the impact parameter of b jets. The algorithm calculates for each track the probability that the track is coming from the primary vertex. These probabilities are then combined together as a likelihood function to obtain the probability, that the jet is coming from the primary vertex. Since the tracks with negative impact parameter are mostly coming from the primary vertex, the compatibility of the tracks to the primary vertex can be calibrated from the negative part of the impact parameter significance distribution. The likelihood function is used as the discriminator to classify jets as b jets or non-b jets depending on the asked discriminator value.

5.2.3 Secondary vertex based algorithm

The secondary vertex based b-tagging algorithm [116] combines the information from the reconstructed secondary vertex as well as the impact point into a likelihood-based discriminator. The algorithm starts by reconstructing the secondary vertex with the trimmed Kalman filter method described in Section 4.3.3.2 and the flight path, i.e. the distance between the primary and secondary vertices, of the assumed b hadron. The algorithm then requires the reconstructed flight path to be greater than 100 μm and less than 2.5 cm, and the transverse flight path significance to be greater than three.

After the cuts on the flight path, the algorithm calculates the invariant mass of the b hadron from the charged tracks, which are associated to the secondary vertex, and requires the invariant mass to be less than 6.5 GeV/c^2. Finally, the secondary vertex is required to be incompatible with 'V^0' decays by rejecting vertices composed of oppositely charged tracks, which have an invariant mass within a 50 MeV/c^2 window of the $\mathrm{K}^0_{\mathrm{S}}$ mass.

To improve the algorithm performance against c jets, the compatibility of the invariant mass calculated from the sum of the tracks to c hadron masses is checked. Finally, the algorithm combines a set of topological and kinematical variables into a likelihood ratio which produces a discriminator value, i.e. a probability, that the jet is or is not a b jet.

5.2.4 Soft lepton identification based algorithm

The b-tagging based on soft electron or muon identification [117] relies on the relatively high ($\sim$20%) semileptonic branching ratio of the b hadron decays. The algorithm starts by extrapolating the tracks,

which are related to the b-jet candidate, to the surface of the ECAL, HCAL, and muon system. Electrons are identified with a neural net approach, which takes as input the variables based on the energy deposition in the ECAL clusters, the ratio of energy deposited in the ECAL and HCAL, and the ratio of momentum carried by the track to the ECAL energy deposition. Muons are identified by requiring a reconstructed muon track to share at least 70% of its hits with a track related to the b-jet candidate. The discriminator value for the b-tagging is then calculated with a neural net, which uses as input the reconstructed lepton momentum, its impact parameter significance, the angular deviation of the lepton from the jet axis, and the ratio of the lepton momentum to the jet energy.

5.2.5 Calibration strategies

The LHC will produce millions of $\mathrm{t\bar{t}}$ events per year, which can be exploited to measure the performance of b-tagging directly from data with b-enriched data samples [118]. The systematic uncertainty of the sample purity determination is dominated by the MC modeling of initial and final state gluon radiation and the uncertainty in the signal and background cross-sections.

The b-tagging efficiency ϵ_{b} can be measured from data with

$$\epsilon_{\mathrm{b}} = \frac{x_{\mathrm{tag}} - \epsilon_0(1 - x_{\mathrm{b}})}{x_{\mathrm{b}}}, \tag{5.5}$$

where x_{tag} is the measured fraction of jets tagged by the b tagging algorithm, ϵ_0 is the mistagging rate for non-b jets in the sample, and x_{b} is the b-jet purity in the sample [52].

Recently, methods in addition to those described above have been implemented from other HEP experiments to measure the b tagging efficiency from data containing muons [119].

5.3 Missing $\mathrm{E_T}$ measurement

The large pseudo-rapidity coverage of the CMS calorimetry systems enable a rather precise momentum and energy conservation in the plane transverse to the beam axis. The conservation of momentum and energy enables the detection of neutrinos, which do not interact with the detector material, as a presence of missing transverse energy, $\mathrm{E_T^{miss}}$. Because of the requirement of complete calorimetry coverage in η,ϕ space, the measurement of $\mathrm{E_T^{miss}}$ is sensitive to detector malfunctions and particles hitting poorly instrumented regions of the detector. Also the detector noise, pile-up events, and underlying event may spoil the $\mathrm{E_T^{miss}}$ measurement resolution.

The $\mathrm{E_T^{miss}}$ is calculated as a vector sum with

$$\vec{\mathrm{E}}_{\mathrm{T}}^{\mathrm{miss}} = \sum_{\mathrm{n}} \left(\mathrm{E_n} \sin\theta_{\mathrm{n}} \cos\phi_{\mathrm{n}} \hat{\mathrm{i}} + \mathrm{E_n} \sin\theta_{\mathrm{n}} \sin\phi_{\mathrm{n}} \hat{\mathrm{j}} \right) = \mathrm{E_x^{miss}}\hat{\mathrm{i}} + \mathrm{E_y^{miss}}\hat{\mathrm{j}}, \tag{5.6}$$

where the sum runs over all calorimeter energy towers. The corrections to the $\mathrm{E_T^{miss}}$ measurement arising from muons, the jet energy scale, the underlying event, and pile-up events are described in Section 5.3.1.

The $\mathrm{E_T^{miss}}$ resolution can be expressed as general parametrization by

$$\sigma(\mathrm{E_T^{miss}}) = \mathrm{A} \oplus \mathrm{B}\sqrt{\sum \mathrm{E_T} - \mathrm{D}} \oplus \mathrm{C}(\sum \mathrm{E_T} - \mathrm{D}), \tag{5.7}$$

where the first term corresponds to electronic noise, pile-up events, and underlying event, the second term is the stochastic term which represents the statistical sampling nature of the energy deposition, and the third term is the constant term which describes the non-linearities, cracks, and dead material. To correct for an offset in the energy resulting from noise and pile-up events, D is substracted from the sum of $\mathrm{E_T}$. With simulation, the parameters have been fitted to A = 1.48$\pm$0.29 GeV, B = 1.04$\pm$0.03 GeV$^{\frac{1}{2}}$, C = 0.023$\pm$0.002, and D = 82$\pm$4 GeV [120]. In this result, the stochastic and constant terms are larger than those given in the physics TDRs, because the calorimeter simulation models the detector performance more realistically.

5.3.1 Corrections to the missing $\mathrm{E_T}$ measurement

The aim of the $\mathrm{E_T^{miss}}$ corrections is to bring the measured $\mathrm{E_T^{miss}}$ value closer to the true $\mathrm{E_T^{miss}}$ by taking into account several effects such as muons, jet energy scale, τ jets, underlying event, pile-up events, etc. [120].

Since the muons are minimum ionizing paricles over a wide range of their momentum, they deposit typically only a few GeV of their energy to the calorimeters. Hence, the muons can generate fake $\mathrm{E_T^{miss}}$. The effect of muons can be corrected by substracting the energy of the muons from the $\mathrm{E_T^{miss}}$ measurement, once the energy deposition of the muons to the calorimeter has been deducted, i.e.

$$\vec{\mathrm{E}}_{\mathrm{T}} = -\sum_{\mathrm{towers}} \vec{\mathrm{E}}_{\mathrm{T}}^{\mathrm{tower}} - \sum_{\mathrm{muons}} \vec{\mathrm{p}}_{\mathrm{T}}^{\mu} + \sum_{\mathrm{deposit\ towers}} \vec{\mathrm{E}}_{\mathrm{T}}^{\mathrm{dep.\ tower}}, \tag{5.8}$$

where the towers correspond to the calorimeter energy towers covering the η,ϕ space. The $\mathrm{E_T^{miss}}$ corrections for the muons are called Type-0 corrections.

After the effect of muons has been corrected, the $\mathrm{E_T^{miss}}$ corrections based on the energy response of the jets, i.e. Type-I corrections, are applied. Typically, MC based corrections similar to those described for τ jets in Section 4.3.1.5 are applied, but the corrections can also be acquired from data [82]. The MC based Type-I $\mathrm{E_T^{miss}}$ corrections are applied with

$$\vec{\mathrm{E}}_{\mathrm{T}}^{\mathrm{corr}} = \vec{\mathrm{E}}_{\mathrm{T}}^{\mathrm{miss}} - \left(\vec{\mathrm{E}}_{\mathrm{T}}^{\mathrm{raw}} + \sum_{\mathrm{jets}} (\vec{\mathrm{E}}_{\mathrm{T}}^{\mathrm{corr.\ jet}} - \vec{\mathrm{E}}_{\mathrm{T}}^{\mathrm{raw\ jet}}) \right), \tag{5.9}$$

where $\mathrm{E}_{\mathrm{Tx(y)}}^{\mathrm{raw}}$ is the sum over the raw calorimeter tower energies from calorimeter towers and the jet

sum is over jets with a reconstructed $\mathrm{E_T^{raw}} > 25$ GeV. The equation (5.9) can be rewritten in the form

$$\vec{\mathrm{E}}_{\mathrm{T}}^{\mathrm{corr}} = \vec{\mathrm{E}}_{\mathrm{T}}^{\mathrm{miss}} - \left(\left[\vec{\mathrm{E}}_{\mathrm{T}}^{\mathrm{raw}} - \sum_{\mathrm{jets}} \vec{\mathrm{E}}_{\mathrm{T}}^{\mathrm{raw\ jet}} \right]_{\mathrm{low}\ \mathrm{E_T}} + \left[\sum_{\mathrm{jets}} \vec{\mathrm{E}}_{\mathrm{T}}^{\mathrm{corr.\ jet}} \right]_{\mathrm{high}\ \mathrm{E_T}} \right) \tag{5.10}$$

which consists of low ($\mathrm{E_T} < 50$ GeV) and the high $\mathrm{E_T}$ ($\mathrm{E_T} > 50$ GeV) parts, which is convenient for the evaluation of systematic uncertainties of the $\mathrm{E_T^{miss}}$ measurement.

The $\mathrm{E_T^{miss}}$ corrections for the effects of underlying event and pile-up events, i.e. Type-II corrections, are yet to be implemented in the reconstruction software and have hence not been applied in the analyses covered in this thesis.

5.4 Top quark reconstruction

The identification of top quark(s) can be applied to many physics channels, such as $\mathrm{t\bar{t}}$ or single top quark production. It is also important in the charged Higgs boson studies, because the production mode $\mathrm{gg} \rightarrow \mathrm{tbH}^{\pm}$ features the production of an associated top quark, whose identification can be used as a constraint against backgrounds.

The top quarks decay almost exclusively via $\mathrm{t} \rightarrow \mathrm{Wb}$ producing hence in addition to the b jet either two jets or a lepton and its neutrino coming from the W decay. Since the W bosons decay most of the time to two jets and since the presence of the neutrino would degrade the $\mathrm{E_T^{miss}}$ measurement in charged Higgs boson studies, only fully hadronic final states are considered in the following for the top quark reconstruction. With this selection, the signature of the top quark decay is three jets, of which one is a b jet, and two of these jets should have an invariant mass close to the W mass.

Typically, an event contains more jets than the three jets needed for the top quark reconstruction. To find the jets, which belong to the top quark decay, constraints based on four-momentum conservation are applied in the form of a kinematic fit. The most general and precise kinematic fit formalism is based on non-linear least squares fit with Lagrange multipliers [118]. This method can be approximated by minimizing the function

$$\chi^2 = \left(\frac{\mathrm{m_{jj}} - \mathrm{m_W}}{\sigma_{\mathrm{W}}} \right)^2 + \left(\frac{\mathrm{m_{jjj}} - \mathrm{m_{top}}}{\sigma_{\mathrm{top}}} \right)^2, \tag{5.11}$$

where $\mathrm{m_{jj}}$ and $\mathrm{m_{jjj}}$ are the invariant masses of a certain combination of two and three jets in the event and σ_{W} and σ_{top} are the Gaussian widths of the reconstructed true W and top quark mass distributions. The correct combination of jets, which come from the top and W decays, is found by minimizing the χ^2 value. The number of combinations can be reduced by requiring, that one of the three jets entering the calculation of the invariant top quark mass is tagged as a b jet. Ambiguous jet combinations are rejected by requiring the χ^2 value to exceed a small value such as 0.1.

6 Search for heavy neutral Higgs bosons in the $\mathrm{H, A} \rightarrow \tau\tau \rightarrow \mathrm{hadr.}$ channel

Observation of a heavy neutral scalar boson, which preferentially couples to b quarks and τ leptons, is an important signature of the MSSM Higgs sector as described in Section 2.2. In the MSSM, at large values of $\tan\beta$, the heavy neutral Higgs bosons H and A are produced in association with b quarks in the $\mathrm{gg} \rightarrow \mathrm{b\bar{b}H(A)}$ process. While the dominant decay channel $\mathrm{H, A} \rightarrow \mathrm{b\bar{b}}$ is overwhelmed by the QCD multi-jet background in proton colliders, the next dominant decay channel, $\mathrm{H, A} \rightarrow \tau\tau$ is accessible. The production cross-section of the $\mathrm{H, A} \rightarrow \tau\tau$ channel is effectively proportional to $\tan^2\beta$ and therefore the $\mathrm{H, A} \rightarrow \tau\tau$ production rate is suitable as an input to global fits to determine the MSSM parameters.

In an earlier study in Ref. [121], it was shown that for high values of $\tan\beta$, the Higgs boson can be discovered in the $\mathrm{H, A} \rightarrow \tau\tau$ channel with 30-60 fb^{-1} of integrated luminosity at CMS, if it exists. The study [122] presented in this section uses improved Monte Carlo samples corresponding to an instantaneous luminosity of $\mathcal{L} = 2\times10^{33}$ cm^{-2}s^{-1}, with full simulation of the CMS detector with GEANT4 [65] and realistic reconstruction techniques for signal and various background samples, including a large QCD multi-jet event data set. The $\mathrm{H, A} \rightarrow \tau\tau$ channel provides an excellent benchmark for the b- and τ-jet identification, as well as for the jet and the missing $\mathrm{E_T}$ reconstruction, which is why it was selected as one of the detailed studies of the physics TDR [25].

In the following, the event samples used for the analysis are described in Section 6.1, the kinematical properties of the signal are explained in Section 6.2, the trigger and offline selections are presented in Section 6.3 and 6.4, respectively, and the results of the offline event selection are shown in Section 6.5. The systematic effects affecting the results and methods used to measure the background from data are discussed in Section 6.6. The discovery reach in the $\mathrm{m_A}$,$\tan\beta$ plane is presented in Section 6.7. Finally, recent developments after the publication of the physics TDRs are described in Section 6.8.

6.1 Event sample generation

The signal events were generated with PYTHIA [69] process $\mathrm{gg} \rightarrow \mathrm{b\bar{b}H(A)}$. The Higgs boson $\mathrm{m_A}$ masses of 200, 500, and 800 GeV/c^2 were simulated and the Higgs boson was forced to decay into two τ jets. The cross-sections and branching ratios of the signal samples were calculated with the FeynHiggs program version 2.3.1 [40, 42, 123]. No preselections were applied for the signal events.

The backgrounds considered were the QCD multi-jet events (for $\tau\tau \rightarrow$ jj mode), $\mathrm{t\bar{t}}$, $\mathrm{b\bar{b}}$, Drell-Yan production of Z/γ^*, W+jet, Wt, and $\tau\tau\mathrm{b\bar{b}}$. These processes were generated with PYTHIA, except for $\tau\tau\mathrm{b\bar{b}}$ which was generated with CompHEP [124]. The QCD multi-jet background was simulated in four bins of transverse momentum of the jets in the rest frame of the hard interaction, $\mathrm{\hat{p}_T}$, with

$50 < \hat{\mathrm{p}}_{\mathrm{T}} <$80 GeV/$c$, $80 < \hat{\mathrm{p}}_{\mathrm{T}} <$120 GeV/$c$, $120 < \hat{\mathrm{p}}_{\mathrm{T}} <$170 GeV/$c$, and $\hat{\mathrm{p}}_{\mathrm{T}} >$170 GeV/$c$.

In the $\mathrm{t\bar{t}}$, Wt and W+jet backgrounds, the W was forced to decay to $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$, and the Z/γ^* Drell-Yan production was required to decay to $\mathrm{Z}/\gamma^* \rightarrow \tau\tau$. In these τ decays, the TAUOLA package [74] was used for τ-lepton decay into all possible decay modes. The W+jet background was generated with PYTHIA processes $\mathrm{q_i\bar{q}_j} \rightarrow \mathrm{gW}$ and $\mathrm{q_ig} \rightarrow \mathrm{q_jW}$ and with jet $\hat{\mathrm{p}}_{\mathrm{T}} > 65$ GeV/c for the transverse momentum defined in the rest frame of the hard interaction. The Z/γ^* generation was split into bins of generated di-τ-lepton mass with $80 < \mathrm{m}_{\tau\tau} <$ 130 GeV/c^2, $130 < \mathrm{m}_{\tau\tau} <$ 300 GeV/c^2, and $\mathrm{m}_{\tau\tau} >$300 GeV/c^2. The $\tau\tau\mathrm{b\bar{b}}$ events were produced in bins of generated di-τ-lepton mass with $60 < \mathrm{m}_{\tau\tau} <$ 100 GeV/c^2 and $\mathrm{m}_{\tau\tau} >$ 100 GeV/c^2. The $\tau\tau\mathrm{b\bar{b}}$ CompHEP events were propagated to PYTHIA for the hadronization with τ lepton allowed to decay into all possible decay modes.

Since the background cross-sections are large compared to the signal cross-sections, loose kinematical pre-selection cuts were applied to all background samples except for the $\tau\tau\mathrm{b\bar{b}}$ process in order to reduce the size of the produced event samples. The preselection cuts were based on jet $\hat{\mathrm{p}}_{\mathrm{T}}$ values and on the requirement to have two "τ-like" jets at the level of generation. The generator level jets were reconstructed with the PYTHIA PYCELL routine with a cone size 0.5. A jet was tagged as "τ-like", if it had $\mathrm{E}_{\mathrm{T}}^{\mathrm{MC}} >$ 50 GeV, $|\eta^{\mathrm{MC}}| < 2.4$ and if the transverse momentum of the leading stable charged particle in the jet satisfied the condition $\mathrm{p}_{\mathrm{T}}^{\mathrm{MC}} >$ 30 GeV/c. These cuts are softer than the τ-jet selection cuts, which were applied at the trigger and offline analysis stages. For the Z/γ^* background, no preselection was applied on the leading track transverse momentum.

The response of the CMS detector was simulated with the CMSIM [79] and OSCAR [78] packages with pile-up events corresponding to the luminosity of $\mathcal{L} = 2\times10^{33}$ cm^{-2}s^{-1}. The events were reconstructed with the ORCA [80] software version 8.7.4. The cross-sections, number of generated events, and the corresponding generated luminosity are summarized in Table 6.1 for the background samples. For the samples with kinematic preselection, the corresponding selection efficiencies, $\epsilon_{\mathrm{presel.}}$, are also shown.

6.2 Signal kinematics

The true τ jet transverse energy, $\mathrm{E}_{\mathrm{T}}^{\mathrm{MC}}$, i.e. the transverse energy of the MC matched τ jet decay products excluding the neutrino, is shown in Fig. 6.1 for the different $\mathrm{m_A}$ values. The average transverse energy of the τ jets was found to increase with the Higgs boson mass. The transverse energy distributions of the τ jets motivate the use of asymmetrical cuts on the $\mathrm{E_T}$ of the two τ jets in the offline selections, which will be explained in Section 6.4.2.

The true pseudo-rapidity of the τ jets is shown in Fig. 6.2 for three $\mathrm{m_A}$ values. It can be seen, that the τ jets are produced as central jets, which are mostly found within the tracker acceptance of $|\eta| < 2.5$.

Samples	$\sigma \times \mathrm{BR}$ (fb)	$\mathcal{L}_{\mathrm{gen.}}$ (fb^{-1})	$\epsilon_{\mathrm{presel.}}$	$\mathrm{N}_{gen.}$
QCD, $\hat{\mathrm{p}}_{\mathrm{T}}$ = 50-80 GeV/c	2.08×10^{10}	0.020	2.44×10^{-4}	100K
QCD, $\hat{\mathrm{p}}_{\mathrm{T}}$ = 80-120 GeV/c	2.94×10^{9}	0.012	5.77×10^{-3}	200K
QCD, $\hat{\mathrm{p}}_{\mathrm{T}}$ = 120-170 GeV/c	5.03×10^{8}	0.009	4.19×10^{-2}	200K
QCD, $\hat{\mathrm{p}}_{\mathrm{T}} > 170$ GeV/c	1.33×10^{8}	0.008	2.12×10^{-1}	1000K
$\mathrm{t\bar{t}}$, $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$	5.76×10^{3}	285	4.88×10^{-2}	80K
Wt, $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$	7.10×10^{2}	3053	1.38×10^{-2}	30K
W+jets, $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$	5.74×10^{5}	32	2.16×10^{-2}	400K
$\mathrm{Z}/\gamma^{*} \rightarrow \tau\tau$, $\mathrm{m}_{\tau\tau}$ = 80-130 GeV/c^2	1.57×10^{6}	4.3	1.90×10^{-2}	128K
$\mathrm{Z}/\gamma^{*} \rightarrow \tau\tau$, $\mathrm{m}_{\tau\tau}$ = 130-300 GeV/c^2	1.24×10^{4}	59	9.53×10^{-2}	70K
$\mathrm{Z}/\gamma^{*} \rightarrow \tau\tau$, $\mathrm{m}_{\tau\tau} > 300$ GeV/c^2	6.22×10^{2}	299	3.23×10^{-1}	60K
$\tau\tau\mathrm{b\bar{b}}$, $\mathrm{m}_{\tau\tau}$ = 60-100 GeV/c^2	2.61×10^{4}	11	1	290K
$\tau\tau\mathrm{b\bar{b}}$, $\mathrm{m}_{\tau\tau} > 100$ GeV/c^2	1.05×10^{3}	95	1	100K

Table 6.1: The cross-sections, number of generated events, and the corresponding generated luminosity for the background samples. For the samples with kinematic preselection, the corresponding selection efficiencies, $\epsilon_{\mathrm{presel.}}$, are also shown.

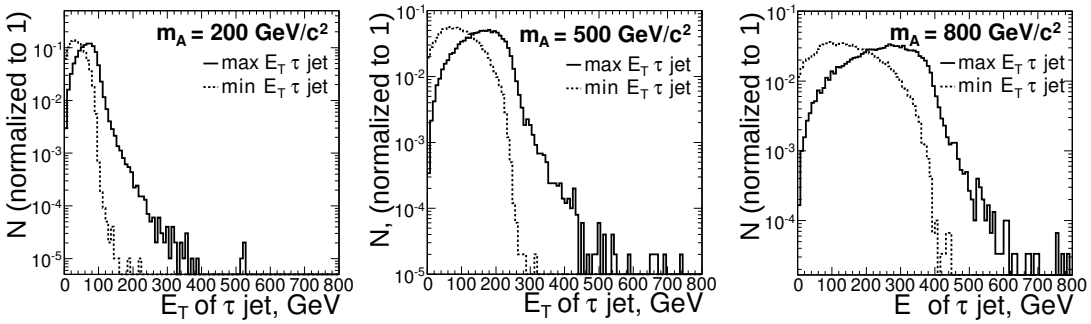

Figure 6.1: The distribution of the true $\mathrm{E_T}$ of the τ jets with the higher (solid histogram) and lower (dashed histogram) $\mathrm{E_T^{MC}}$ for the signal samples with $\mathrm{m_A}$ = 200 GeV/c^2 (left), $\mathrm{m_A}$ = 500 GeV/c^2 (middle), and $\mathrm{m_A}$ = 800 GeV/c^2 (right).

6.3 First and high level trigger selections

The $\tau\tau \rightarrow$ jj final state was triggered by the single or double Level-1 τ trigger described in Section 5.1.3. The trigger thresholds for the Level-1 jets were set to 93 GeV for the single and 66 GeV for the double τ trigger [111]. The Level-1 trigger was followed by the double τ high level trigger described in Section 5.1.4. Of the two selection strategies available at the high level trigger, the Calo+Pxl selection was used. In this approach, the electromagnetic isolation is applied to the τ-jet candidate with the highest transverse energy followed by charged track isolation based on information of the pixel detector alone. The Calo+Pxl selection was applied to both τ-jet candidates.

In the following, the high-level trigger performance is described in more detail for the HLT τ-jet candidate with the highest transverse energy and for both HLT τ-jet candidates [125]. In addition to the $\mathrm{b\bar{b}H(A)}$ signal and the QCD multi-jet background samples, a sample containing back-to-back double τ jets (pure τ) described in Section 5.1.1 was used. The efficiencies in the following results have been calculated with

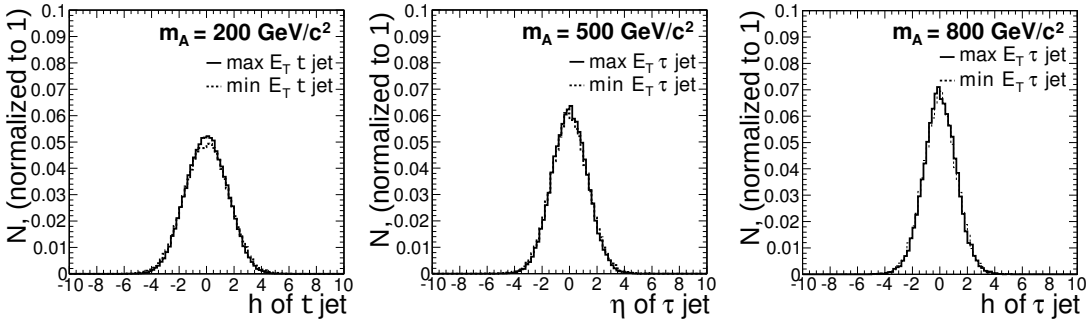

Figure 6.2: The distribution of the jet η of the τ jets with the higher (solid histogram) and lower (dashed histogram) $\mathrm{E_T^{MC}}$ for the signal samples with $\mathrm{m_A}$ = 200 GeV/c^2 (left), $\mathrm{m_A}$ = 500 GeV/c^2 (middle), and $\mathrm{m_A}$ = 800 GeV/c^2 (right).

respect to the events which have passed the Level-1 single or double τ trigger.

The efficiency for passing the Calo+Pxl high level trigger on the Level-1 τ-jets with the highest energy is shown in Fig. 6.3 as a function of the isolation cone size ($\mathrm{R_i}$) of the charged track isolation algorithm for the signal and background samples. For the pure τ events (2-tau in Fig. 6.3), the high level trigger efficiency was found to be $\sim$86% independent of the isolation cone size. When pile-up events corresponding to the LHC low luminosity scenario was added to the pure τ events (2-tau PU in Fig. 6.3), the efficiency was decreased by $\sim$4% as a result of the inefficiencies in the high level trigger jet reconstruction and pixel detector. For τ jets from the $\mathrm{b\bar{b}H}$ signal samples, the efficiency for passing the Calo+Pxl trigger was found not to depend on the Higgs boson mass as can be seen in Fig. 6.3. The requirement of a reconstructed primary vertex was observed to decrease the efficiency by few percent units. For QCD multi-jet events, the efficiency for passing the Calo+Pxl trigger can be seen to fall until $\mathrm{R_i}$ of 0.35-0.40 after which the gain in the rejection slows down. A detailed breakdown of the sources of inefficiencies, why events fail the high level trigger trigger for the τ-jet candidate with the highest transverse energy, is presented in Table 6.2 for an isolation cone size of $\mathrm{R_i}$ = 0.35.

Event type	No HLT jets	No tracks in jet cone	No leading track	Not isolated	No primary vertex
pure-tau	2.9%	9.5%	1.2%	0.4%	-
pure-tau-PU	5.0%	10.0%	2.2%	0.8%	-
$\mathrm{b\bar{b}H}$(200)	9.4%	7.5%	5.1%	11%	3.4%
$\mathrm{b\bar{b}H}$(500)	9.0%	7.9%	3.2%	9.3%	4.4%
$\mathrm{b\bar{b}H}$(800)	9.3%	8.7%	3.2%	9.9%	4.8%
QCD, $\mathrm{\hat{p}_T}$ = 50-80 GeV/c	49%	1.7%	6.1%	36%	0.3%
QCD, $\mathrm{\hat{p}_T}$ = 80-120 GeV/c	64%	0.8%	3.1%	27%	0.2%
QCD, $\mathrm{\hat{p}_T}$ = 120-170 GeV/c	75%	0.5%	1.6%	20%	0.2%

Table 6.2: Sources of a negative HLT response for different event types with $\mathrm{R_i}$ = 0.35 for the Level-1 τ jet with the highest energy.

The efficiency of the Calo+Pxl high level trigger for two Level-1 τ jets is shown in Fig. 6.4 on linear and

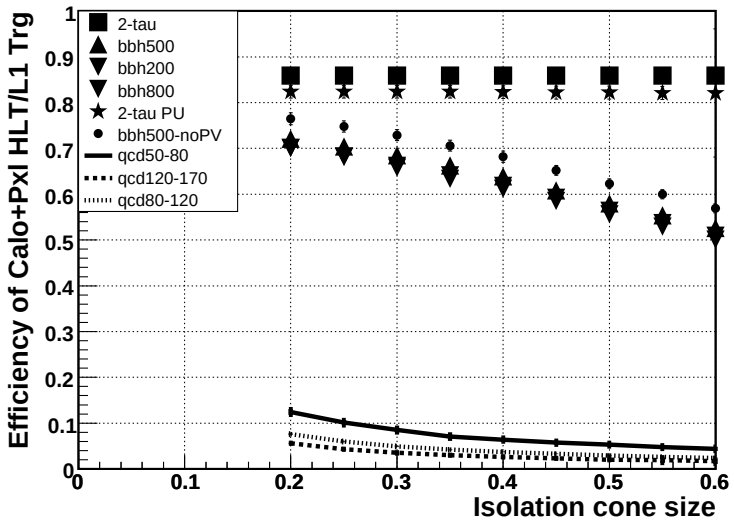

Figure 6.3: Efficiency of the Calo+Pxl τ high level trigger for the Level-1 τ-jets with the highest transverse energy as a function of the isolation cone size of the charged track isolation algorithm.

logarithmic scales as a function of the isolation cone size ($\mathrm{R_i}$) of the charged track isolation algorithm for the signal and background samples. As for the single τ jet case, the efficiency of the pure τ sample is observed to remain independent of the isolation cone size. The efficiency of the signal samples can be seen to diminish by $\sim$20% units as a result of applying the pixel track isolation also to the second jet. The same requirement was, however, found to reduce the QCD multi-jet events by an additional factor of $\sim$10.

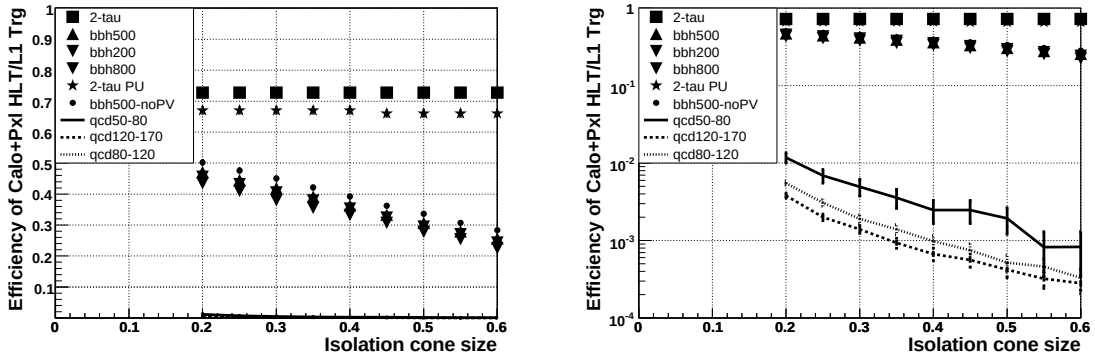

Figure 6.4: Efficiency of the Calo+Pxl HLT for both Level-1 τ jets versus the isolation cone size $\mathrm{R_i}$ on linear (left plot) and logarithmic (right plot) scale.

An attempt was made to vary the cut on the $\mathrm{E_T}$ threshold of the HLT jet to gain additional background rejection. It was observed, that a notable reduction of background would require applying a threshold of 50-60 GeV on the jet $\mathrm{E_T}$. Such a threshold was, however, found to reduce the signal samples too much, especially for the light Higgs boson ($\mathrm{m_A}$ = 200 GeV/c^2).

The performance of the Calo+Pxl high level trigger algorithm was evaluated in more detail for the $\mathrm{b\bar{b}H}$(500) sample. The efficiency for reconstructing the HLT τ jets and the overall efficiency of the

Calo+Pxl high level trigger with respect to the Level-1 trigger are shown in Fig. 6.5 as a function of the simulated τ lepton transverse momentum. The efficiency for reconstructing the HLT τ jets was found to increase rapidly and reaches an almost 100% plateau for simulated τ jet $\mathrm{p_T}$ in excess of 60 GeV/c. It was also established from Fig. 6.5, that $\sim$90% of the HLT τ jets were well matched to the direction of the simulated τ jets, when a matching cone of 0.5 was used. The efficiency of the Calo+Pxl high level trigger algorithm was observed to increase with the $\mathrm{p_T}$ of the simulated τ jet reaching a plateau of $\sim$65% for $\mathrm{p_T} > 60$ GeV/c. Practically all τ jet candidates, which passed the Calo+Pxl high level trigger algorithm, were observed to match to a simulated τ jet.

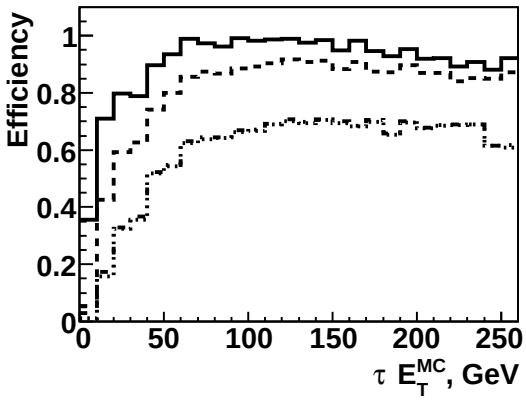

Figure 6.5: The efficiency of the HLT τ jet reconstruction (solid line), matching of the HLT τ jets to the simulated τ jets (dashed line), Calo+Pxl trigger (dashed-dotted line), and matching of the Calo+Pxl jets to the simulated τ jets (dotted line) as a function of the simulated τ lepton transverse momentum for the $\mathrm{b\bar{b}H}$(500) sample.

The efficiency of the Calo+Pxl trigger is shown in Fig. 6.6 as a function of the simulated τ lepton transverse momentum and pseudo-rapidity for pure τ jets from the $\mathrm{b\bar{b}H}$ samples with $\mathrm{m_A}$ = 200 and 500 GeV/c^2. It can be seen in Fig. 6.6, that the plateau of the Calo+Pxl trigger efficiency is reached at $\sim$60 GeV/c for simulated τ lepton transverse energy for all simulated samples. The Calo+Pxl trigger efficiency was observed to remain flat as a function of the simulated τ lepton pseudo-rapidity up to η = 2.1 after which efficiency was lost as a result of the loss of coverage of the pixel detector. The Calo+Pxl trigger efficiency was found to be similar for the three Higgs boson masses.

6.4 Offline event reconstruction and selections

If an event passed the Level-1 and HLT single or double τ triggers, the τ-jet candidates and the other jets were reconstructed with information available from all subdetectors. Then, a set of offline selections were applied to it.

The first step in the offline analysis was the τ-jet identification which was applied to both τ-jet candidates. It was especially aimed against the QCD multi-jet (including $\mathrm{b\bar{b}}$) and W+jet backgrounds. The τ-jet

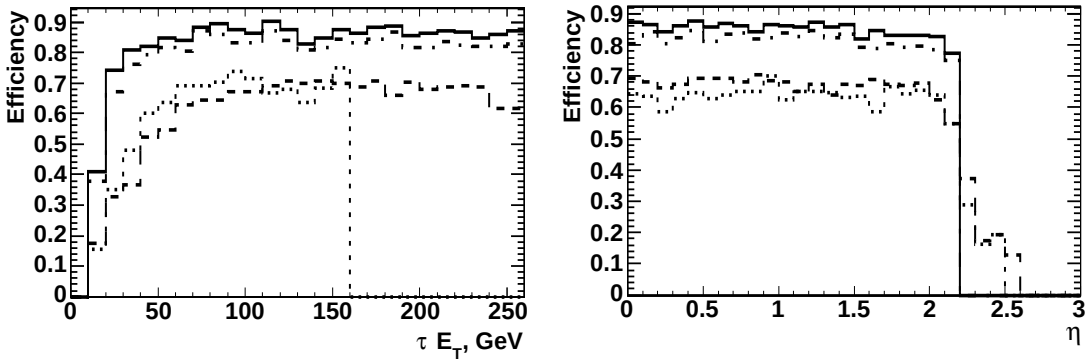

Figure 6.6: Comparison of the Calo+Pxl high level trigger efficiency as a function of transverse momentum (left) and pseudo-rapidity (right) of the reconstructed τ jets for the double τ sample without pile-up events (solid line), double τ sample with pile-up events (dashed-dotted line), and the $\mathrm{b\bar{b}H}$(500) (dashed line) and $\mathrm{b\bar{b}H}$(200) (dotted line) event samples.

identification included setting a cut on the leading track $\mathrm{p_T}$, requiring charged track isolation with more stringent cuts than at the high level trigger based on fully reconstructed tracks and demanding one or three tracks to be found in the signal cone. Additionally, the total charge of the tracks of the two τ jets was required to be opposite. The τ-jet identification methods based on the impact parameter of the leading track and on the τ-jet mass reconstruction were tried, but they were found not to improve the significance of the signal.

Since the associated $\mathrm{b\bar{b}H, A}$ production dominates the Higgs boson production at high values of $\tan\beta$, it is natural to apply b-jet tagging, which suppresses the Drell-Yan $\tau\tau$ production and which suppresses further the QCD multi-jet and W+jet backgrounds. Since the transverse energy of the associated b jets is relatively small in the signal, the tagging of only a single b jet was required in the analysis. Finally, a veto was placed on events with jets in addition to the identified τ and the b jets in order to reduce the $\mathrm{t\bar{t}}$ and Wt backgrounds.

In the following subsections, the jet reconstruction, τ-jet identification, the missing $\mathrm{E_T}$ reconstruction, and the b-jet tagging are described in more detail.

6.4.1 Jet and track reconstruction

The two τ-jet candidates were reconstructed with the iterative cone algorithm. The calorimeter towers within a cone of 0.8 in η,ϕ space around the HLT jets were used as input for the jet reconstruction. The tower seed threshold was set to 1 GeV. The uncorrected energy, $\mathrm{E_T^{raw}}$, of the calorimeter tower with the highest energy deposition was required to exceed 10 GeV. The efficiency of this requirement was almost 100% for the signal events. The jets were reconstructed with the iterative cone algorithm described in Section 4.3.1.1. The jet cone size of 0.4 was chosen for reconstructing the jets in the signal samples based on the optimization in Section 5.1. The purity, i.e. the probability, that the jet direction matches the generated τ lepton direction within a matching cone of 0.4, was found to be $\sim$100%.

The other calorimeter jets in the event were reconstructed with the iterative cone algorithm with a jet cone size of 0.5. In order to reduce the fake jet contamination from the electronic noise and pile up events [126], the thresholds for the calorimeter tower energy were set to 0.5 and 0.8 GeV for the barrel and endcap towers, respectively, for both all types of jets.

The Monte Carlo based jet energy corrections described in Section 5.1.13.2 were used for the τ jets. For other jets, the Monte Carlo based calibration described in Ref. [126] were used. The corrections were obtained in both cases in the presence of pile-up events corresponding to the LHC low luminosity scenario of $\mathcal{L} = 2\times10^{33}$ cm^{-2}s^{-1}. The τ-jet energy scale, i.e. the ratio $\mathrm{r} = \mathrm{E_T^{reco}}/\mathrm{E_T^{MC}}$, is shown before and after the jet energy corrections in Fig. 6.7 for τ jets from the signal events as a function of simulated τ jet transverse energy and pseudo-rapidity. The loss of efficiency at $|\eta| \sim 1.3$ is caused by the gap between the barrel and endcap parts in the calorimetry system. The ratios $\mathrm{r_{raw}} = \mathrm{E_T^{raw}}/\mathrm{p_T}^{\mathrm{b}}$ and $\mathrm{r_{calibr.}} = \mathrm{E_T^{calibr.}}/\mathrm{p_T}^{\mathrm{b}}$ of the associated b jets are shown in Fig. 6.8 to verify the Monte Carlo jet energy calibration for b jets. The fitted mean value and standard deviation for the solid histogram were found to be 0.97 and $\sigma = 0.16$, respectively.

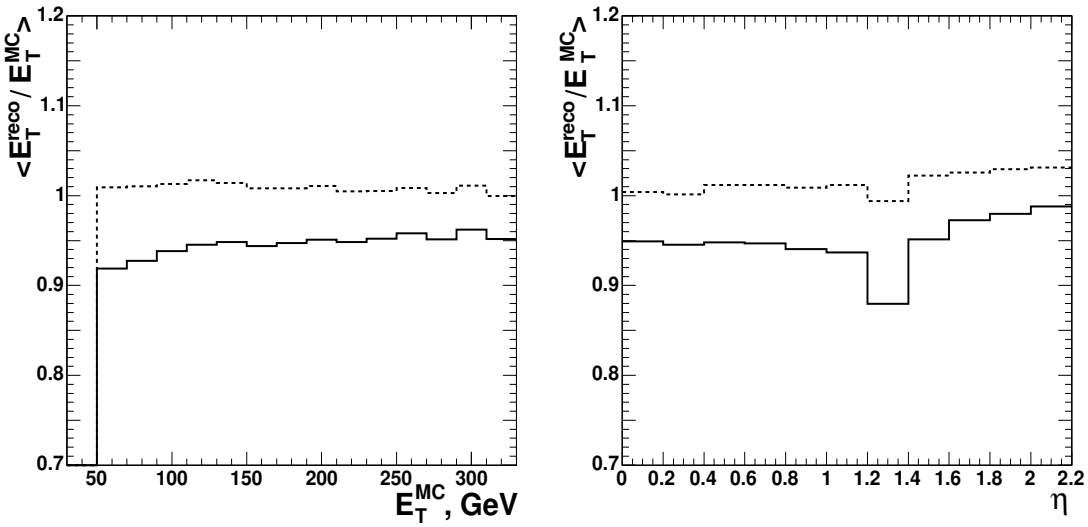

Figure 6.7: The ratio $\mathrm{r} = \mathrm{E_T^{reco}}/\mathrm{E_T^{MC}}$ as a function of the simulated τ jet transverse energy (left) and pseudo-rapidity (right) before (solid histogram) and after (dashed histogram) the τ jet energy corrections were applied for all signal samples.

The tracks in the events were reconstructed with the combinatorial Kalman filter and smoother described in Sections 4.3.2.1 and 4.3.2.2 from track seeds generated with the pixel detector information.

6.4.2 Identification of the τ jets

To identify the τ jets, a threshold was set on the transverse energy of the τ-jet candidates and on the transverse momentum of the leading track in the jets. These requirements were followed by charged track isolation with either one or three tracks in the signal cone. These selections are described in more detail in

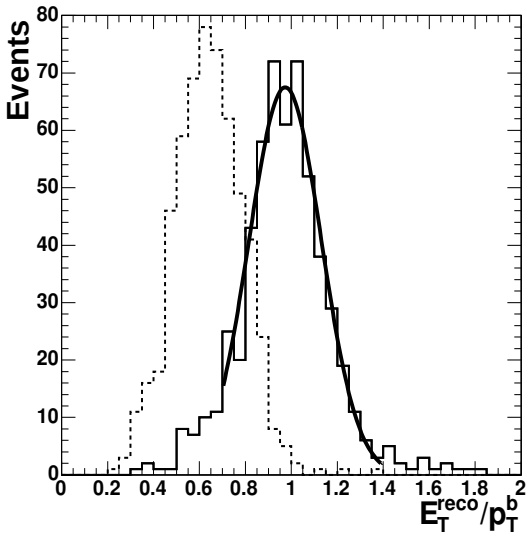

Figure 6.8: The ratios $\mathrm{r_{raw}} = \mathrm{E_T^{raw}}/\mathrm{p_T}^{\mathrm{b}}$ (dashed line) and $\mathrm{r_{calibr.}} = \mathrm{E_T^{calibr.}}/\mathrm{p_T}^{\mathrm{b}}$ (solid line) for the non-τ jet in the event associated with the b quark.

the following. Additionally, the τ-jet identification method based on the impact parameter of the leading track and the reconstruction of the visible τ-jet mass were tried. These methods were, however, found not to improve the signal significance. The rejection power of the visible τ-jet mass was found to be very small after the trigger selections.

The raw transverse energy of the two off-line τ-jet candidates, $\mathrm{E_T^{\tau\,jet}}$, is shown in Figs. 6.9 and 6.10 for the signal and QCD multi-jet background, respectively, after the trigger selections. As expected, the $\mathrm{E_T}$ spectra of the τ-jet candidate with the highest transverse energy (first jet) and the τ-jet candidate with the lowest energy (second jet) are not very different in the QCD multi-jet events, whereas they are different for the signal events. The difference motivates the usage of asymmetrical thresholds of the transverse energy of the first and the second τ-jet candidate.

The uncorrected transverse energy of both jets was required to exceed 50 GeV at the generator level preselection for the QCD multi-jet background. This requirement was found to effectively suppress the QCD multi-jet events with $30 < \mathrm{E_T^{MC}} < 50$ GeV. It should be noted, that these thresholds have already been applied in Fig. 6.10. The threshold on the transverse energy of the first jet was varied depending on the mass of the Higgs boson. It was set to 50, 100, and 150 GeV for the $\mathrm{m_A}$ = 200, 500, and 800 GeV/c^2, respectively. The uncorrected transverse energy of the second jet was required to exceed 50 GeV in accordance with the generator level preselections.

After passing the thresholds on the transverse energy of the τ-jet candidates, the tracks belonging to the τ-jet candidates were examined. The tracks within a cone of 0.5 in η,ϕ space around the τ-jet candidate axis were considered to belong to the jet. Of these tracks, the track with the highest transverse momentum, i.e. the leading track, was required to be within a matching cone of 0.1 around the τ-jet candidate direction. The transverse momentum of the leading track of τ-jet candidates is shown in Fig. 6.11 for the signal and in Fig. 6.12 for the QCD multi-jet background. It can be seen in Fig. 6.12,

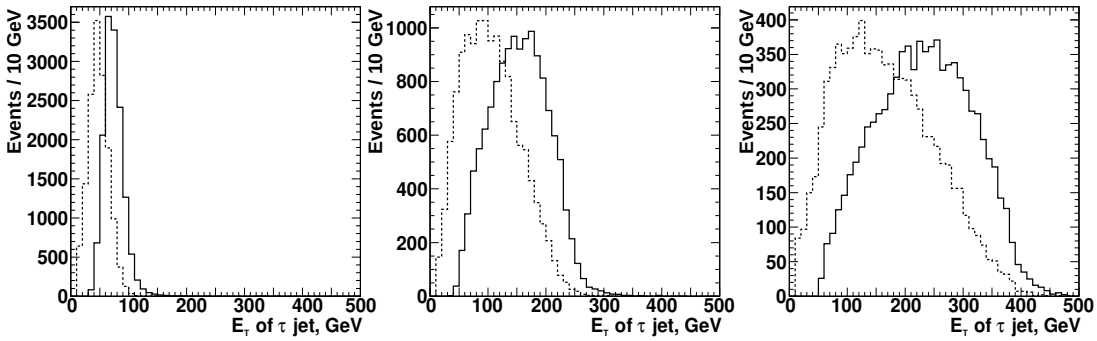

Figure 6.9: The uncorrected transverse energy of the τ-jet candidate with the highest $\mathrm{E_T^{raw}}$ (solid line) and of the τ-jet candidate with the lowest $\mathrm{E_T^{raw}}$ (dashed line) for signal events with $\mathrm{m_A}$=200 GeV/c^2 (left plot), 500 GeV/c^2 (middle plot), and 800 GeV/c^2 (right plot).

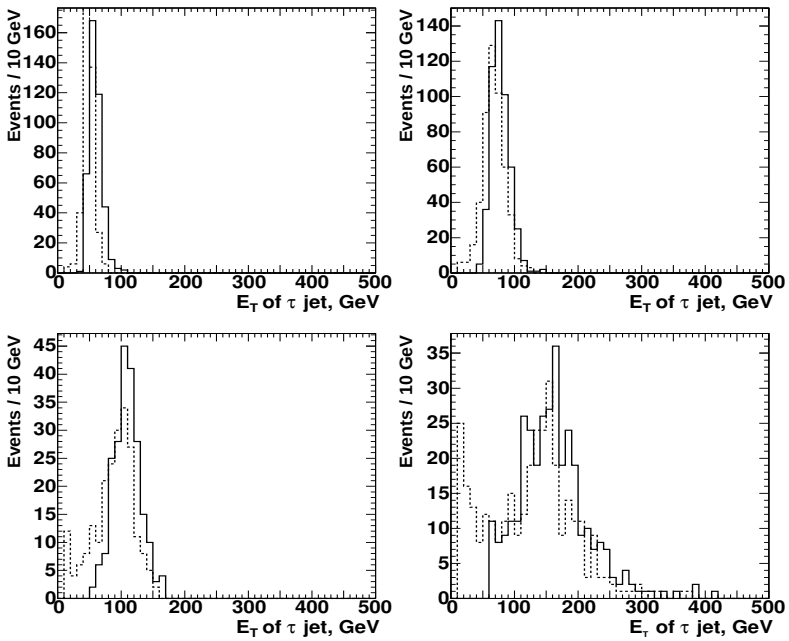

Figure 6.10: The uncorrected transverse energy of the τ-jet candidate with the highest $\mathrm{E_T^{raw}}$ (solid line) and of the τ-jet candidate with the lowest $\mathrm{E_T^{raw}}$ (dashed line) for QCD multi-jet events with $50 < \hat{\mathrm{p}}_\mathrm{T} < 80$ GeV/c (top left), $80 < \hat{\mathrm{p}}_\mathrm{T} < 120$ GeV/c (top right), $120 < \hat{\mathrm{p}}_\mathrm{T} < 170$ GeV/c (bottom left), and $\hat{\mathrm{p}}_\mathrm{T} > 170$ GeV/c (bottom right).

that even though the transverse momentum of the leading track was required to exceed 30 GeV/c at generator level, some tracks were reconstructed with a transverse momentum below this cut value for the QCD multi-jet events. This was found to be caused by events in which the leading track was not reconstructed properly. The transverse momentum of the leading track was required to exceed 35 GeV/c in accordance to Ref. [111].

If the τ-jet candidates satisfied the cuts for the transverse energy of the jet and the transverse momentum of the leading track in the jet, charged track isolation described in Section 5.1.6 was applied to the candidate. Only tracks with transverse momentum exceeding 1.0 GeV/c and with impact parameter

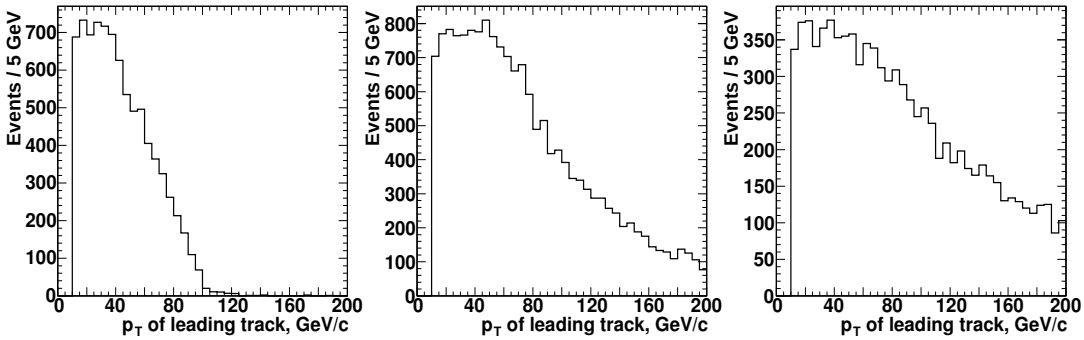

Figure 6.11: The transverse momentum of the leading track in τ-jet candidates for the signal with $\mathrm{m_A}$ = 200 GeV/c^2 (left plot), 500 GeV/c^2 (middle plot), and 800 GeV/c^2 (right plot) after passing the trigger and jet $\mathrm{E_T}$ thresholds.

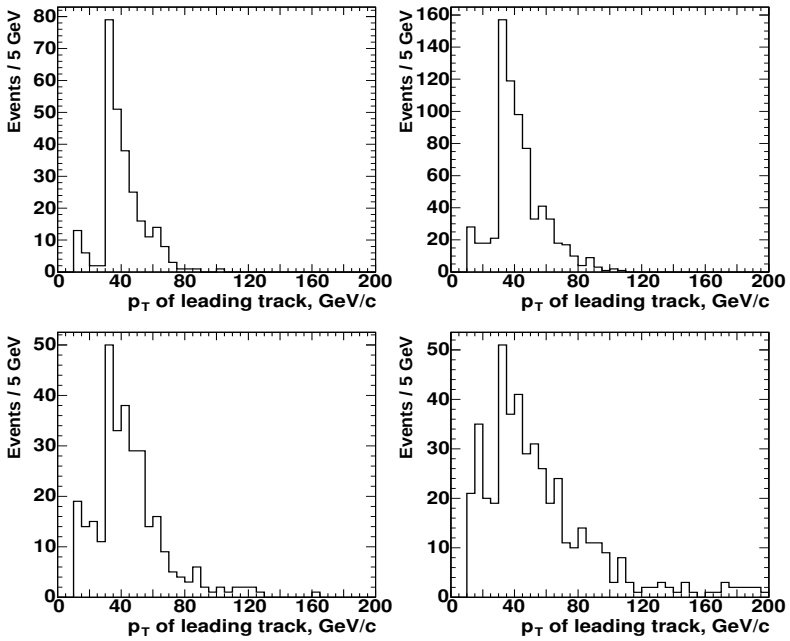

Figure 6.12: The transverse momentum of the leading track of the τ-jet candidate for hadronic jets from QCD multi-jet events with $50 < \hat{\mathrm{p}}_\mathrm{T} < 80$ GeV/c (top left), $80 < \hat{\mathrm{p}}_\mathrm{T} < 120$ GeV/c (top right), $120 < \hat{\mathrm{p}}_\mathrm{T} < 170$ GeV/c (bottom left), and $\hat{\mathrm{p}}_\mathrm{T} > 170$ GeV/c (bottom right) after passing the trigger and jet $\mathrm{E_T}$ thresholds.

along the beam axis within 2 mm from that of the leading track were counted. The sizes of the matching and isolation cones were set to $\mathrm{R_m}$ = 0.1 and $\mathrm{R_i}$ = 0.5, respectively. The tracks were required to have at least eight hits of which at least two in the pixel detector and the normalized χ^2 of the track fit was required to be less than 10.

The τ-jet candidate was labeled as a τ jet, if no tracks matching the above criteria were found within the annulus between the signal and isolation cones and if the number of tracks within the signal cone matched to a preset number. The number of reconstructed tracks in the signal cone of $\mathrm{R_s}$ = 0.04 is

shown in Fig. 6.13 for the signal and in Fig. 6.14 for the QCD multi-jet background. It can be seen, that the selection of only one track in the signal cone could be beneficial. The decrease in the fraction of signal τ jets with three reconstructed tracks in the signal cone as $\mathrm{m_H}$ increases results from the increased collimation of the three tracks with the τ jet $\mathrm{E_T^{MC}}$. To find the optimal selection parameters, the following variations of the size of the signal cone, $\mathrm{R_s}$, and the number of tracks required in the signal cone, $\mathrm{N_{trk}^{signal}}$, were considered:

- one or three tracks in the signal cone, $\mathrm{R_s}$ = 0.07
- one or three tracks in the signal cone, $\mathrm{R_s}$ = 0.04
- one track in the signal cone, $\mathrm{R_s}$ = 0.04

The expected number of events with 60 fb^{-1} of data for the signal of $\mathrm{m_A}$ = 200 GeV/c^2 and $\tan\beta$ = 20 in the $\mathrm{m_h^{max}}$ MSSM scenario [39, 127] and for the QCD multi-jet events with the different combinations of $\mathrm{R_s}$ and $\mathrm{N_{trk}^{signal}}$ values are shown in Table 6.3. The trigger and all the offline event selections described in Sections 6.4.2-6.4.5 were applied to calculate the numbers in the table. It can be seen from Table 6.3, that the choice of the smaller signal cone size reduces the QCD multi-jet background and improves the signal significance. The requirement to have only one track in the signal cone was observed to improve the signal-to-background ratio (S/B) for the signal with $\mathrm{m_A}$ = 200 GeV/c^2, albeit reducing the significance of the discovery. However, for the signal samples with $\mathrm{m_A}$ = 500 and 800 GeV/c^2, the requirement $\mathrm{N_{trk}^{signal}}$ = 1 in the signal cone size of 0.04 was found to yield better signal significance than the $\mathrm{N_{trk}^{signal}}$ = 1 or 3 requirement. Hence only one track was required for the searches of Higgs bosons with $\mathrm{m_A}$ = 500 and 800 GeV/c^2. The signal cone size of $\mathrm{R_s}$ = 0.04 was chosen, since it was found to yield a better separation against the QCD multi-jet background than $\mathrm{R_s}$ = 0.07.

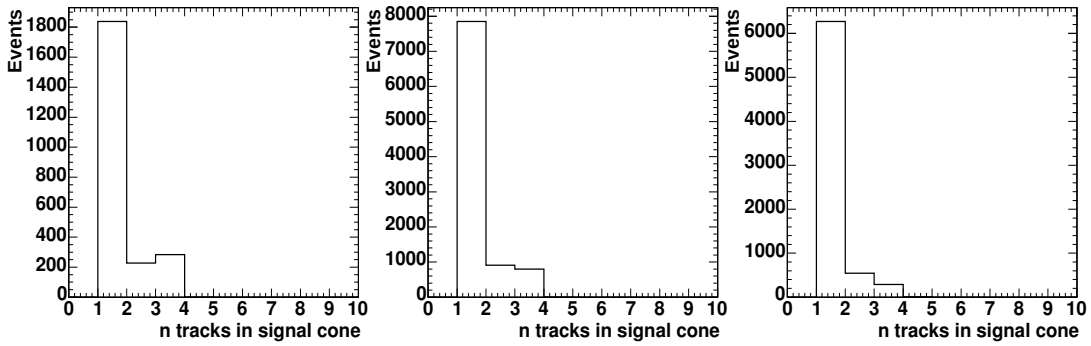

Figure 6.13: The number of the reconstructed tracks in the signal cone of $\mathrm{R_s}$ = 0.04 for the signal with $\mathrm{m_A}$ = 200 GeV/c^2 (left plot), 500 GeV/c^2 (middle plot), and 800 GeV/c^2 (right plot).

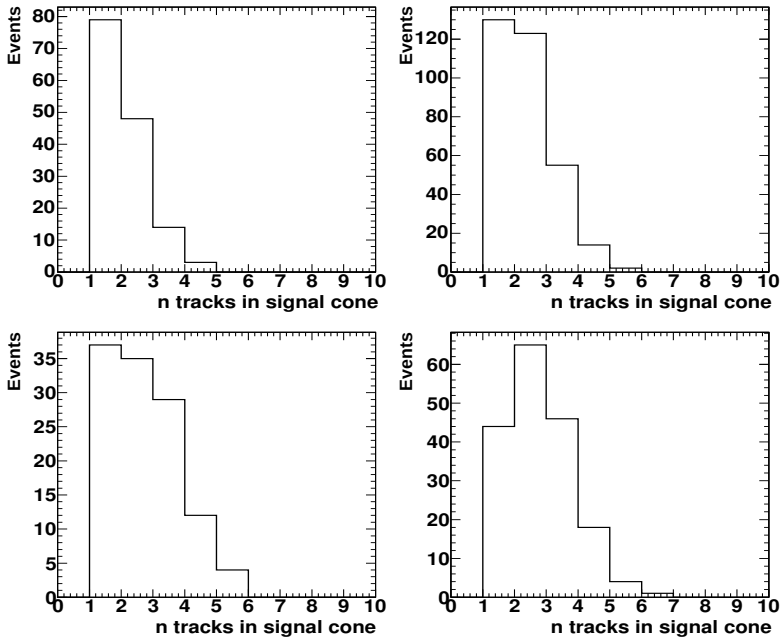

Figure 6.14: The number of the reconstructed tracks in the signal cone of $\mathrm{R_s}$ = 0.04 for QCD multi-jet events with $50 < \hat{\mathrm{p}}_\mathrm{T} < 80$ GeV/c (top left), $80 < \hat{\mathrm{p}}_\mathrm{T} < 120$ GeV/c (top right), $120 < \hat{\mathrm{p}}_\mathrm{T} < 170$ GeV/c (bottom left), and $\hat{\mathrm{p}}_\mathrm{T} > 170$ GeV/c (bottom right).

	Signal	QCD multi-jet background in $\hat{\mathrm{p}}_\mathrm{T}$ bins, GeV/c				$\mathrm{S}/\sqrt{\mathrm{B}}$
	$\mathrm{m_A}$=200 GeV/c^2	50-80	80-120	120-170	>170	
1 or 3 trk, $\mathrm{R_s}$=0.07	67.7	6.9	67.4	77.5	36.0	4.9
1 or 3 trk, $\mathrm{R_s}$=0.04	58.2	4.4	40.6	38.4	20.7	5.7
1 trk, $\mathrm{R_s}$=0.04	37.4	3.1	19.3	21.9	5.6	5.3

Table 6.3: The expected number of events and significance of the discovery with an integrated luminosity of 60 fb^{-1} for signal with $\mathrm{m_A}$ = 200 GeV/c^2 and $\tan\beta$ = 20 in the $\mathrm{m_h^{max}}$-scenario and for QCD multi-jet background in different $\hat{\mathrm{p}}_\mathrm{T}$ bins. The values are evaluated for three sets of $\mathrm{R_s}$ and $\mathrm{N_{trk}^{signal}}$ values.

6.4.3 Identification of associated b jets

The associated Higgs boson production mode $\mathrm{gg} \rightarrow \mathrm{b\bar{b}H(A)}$ provides the opportunity to identify one or two b jets in the events. Figures 6.15 and 6.16 show the transverse momentum and pseudo-rapidity distributions, respectively, of the two b quarks for the signal samples with $\mathrm{m_A}$ = 200, 500, and 800 GeV/c^2. It can be seen, that the b quark with the lowest transverse momentum is very soft and often outside the tracker acceptance. Since the b-tagging algorithms are based on information based on tracks, the presence of only one b-tagged jet was required in the event.

The track counting algorithm described in Section 5.2.1 was used for identifying the b jets. The three-dimensional impact parameter significance of at least three of the tracks associated to the b-jet candidate was required to exceed the b-tagging discriminator value. The resulting b-tagging efficiency

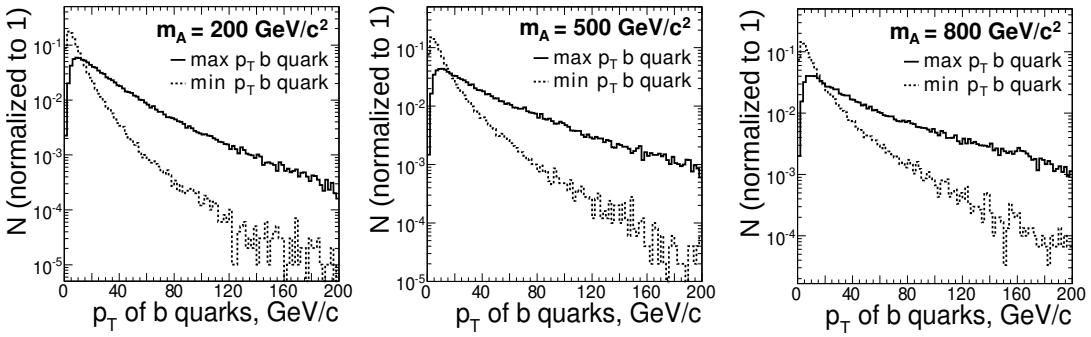

Figure 6.15: The distribution of the transverse momentum for the b quark with the highest $\mathrm{p_T}$ (solid line) and lowest $\mathrm{p_T}$ (dashed line) for the signal sample with $\mathrm{m_A}$ = 200 (left), 500 (middle), and 800 GeV/c^2 (right).

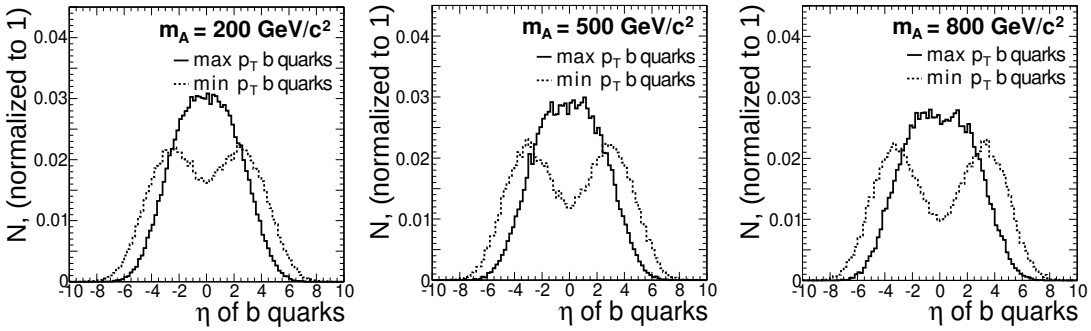

Figure 6.16: The distribution of the pseudo-rapidity for the b quark with the highest $\mathrm{p_T}$ (solid line) and lowest $\mathrm{p_T}$ (dashed line) for the signal sample with $\mathrm{m_A}$ = 200 (left), 500 (middle), and 800 GeV/c^2 (right).

is shown in Fig. 6.17 for the signal with $\mathrm{m_A}$ = 500 GeV/c^2, and for the QCD multi-jet events with 120 $<\hat{\mathrm{p}}_\mathrm{T}<$170 GeV/$c$ as a function of the b-tagging discriminator value. Based on Fig. 6.17, the b-tagging discriminator value was required to exceed 2.0, which is the default value of the algorithm. The b-tagging efficiency is shown in Fig. 6.18 as a function of the $\mathrm{p_T}$ of the b quark for the signal samples. The b-tagging efficiency was observed to increase with the $\mathrm{p_T}$ of the b quark for both the signal and QCD multi-jet events.

6.4.4 Veto on additional central jets

So far, the identification of two τ jets and a third jet tagged as a b jet have been required in the event. The distributions of the calibrated transverse energy $\mathrm{E_T^{calibr.}}$ of this third jet and of third jets associated to b quarks are shown in Fig. 6.19 for the signal with $\mathrm{m_A}$ = 500 GeV/c^2. The simulated momentum of the b quarks, which were associated to the third jets, is shown in the figure as a dotted line. The association was done by requiring the jet and the b quark directions to coincide within a distance of 0.4 in η,ϕ space. Based on Fig. 6.19, the transverse energy of the third jet was required to exceed 20 GeV. Additionally the direction of third jet was required to be within $|\eta| < 2.4$.

Since the second b jet is generally very soft or outside the tracker acceptance, a veto was placed on other

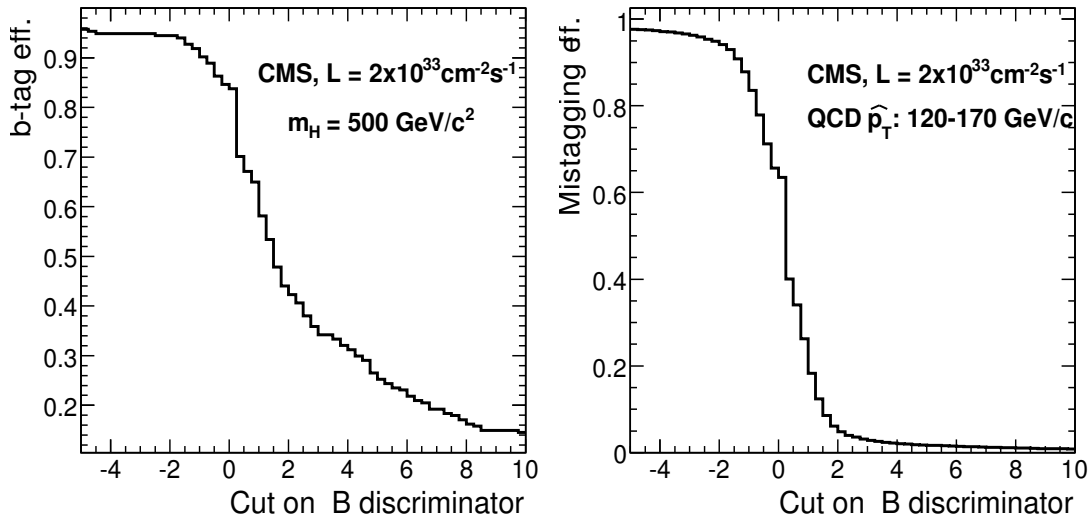

Figure 6.17: The b-tagging efficiency as a function of the lower cut on the b-tagging discriminator for the signal of $\mathrm{m_A}$ = 500 GeV/c^2 (left plot) and the QCD multi-jet background generated with $\hat{\mathrm{p}}_\mathrm{T}$=120-170 GeV/$c$ (right plot).

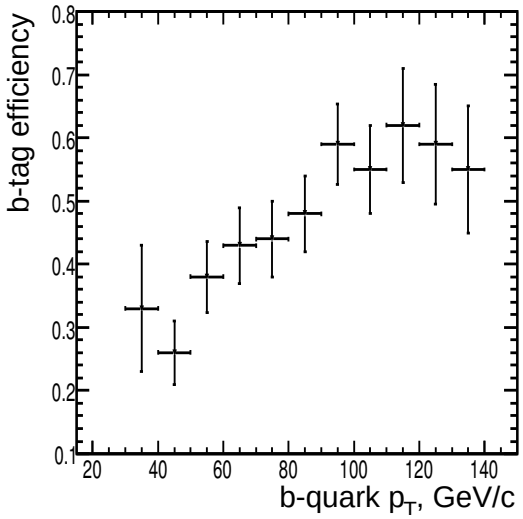

Figure 6.18: The b-tagging efficiency as a function of the $\mathrm{p_T}$ of the b quark for the signal.

central jets than the two τ jets and the third jet to suppress the $\mathrm{t\bar{t}}$ background. If additional jets were found with $\mathrm{E_T} > 20$ GeV and with $|\eta| < 2.4$, the event was rejected.

6.4.5 Missing $\mathrm{E_T}$ and Higgs boson mass reconstruction

In spite of the escaping neutrinos, the Higgs boson mass can be reconstructed in the $\mathrm{H, A} \rightarrow \tau\tau$ channel from the visible τ-jet momenta, i.e. reconstructed τ-jet energy, and the missing transverse energy, $\mathrm{E_T^{miss}}$, with the collinearity approximation ($\mathrm{p}_\tau \gg \mathrm{m}_\tau$) for neutrinos from highly boosted τ's. In the collinearity

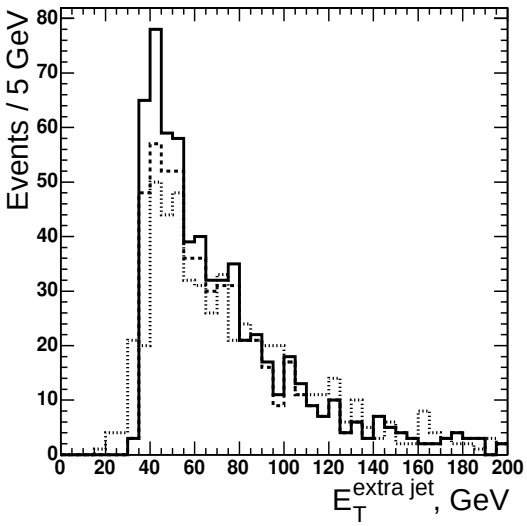

Figure 6.19: the Monte Carlo calibrated energy $\mathrm{E_T^{calibr.}}$ for the third jet (solid line) and for the third jets associated with b quark (dashed line) in the event for the signal of $\mathrm{m_A}$ = 500 GeV/c^2. The simulated momentum distribution of the b quark associated with the third jet is shown as dotted line.

approximation, the τ neutrinos are assumed to be collinear with the τ jets, and thus the $\mathrm{E_T^{miss}}$ vector is divided into components in the two τ-jet directions.

The measurement of the $\mathrm{E_T^{miss}}$ is affected by the non-linear calorimeter response. Earlier studies in Refs. [110, 128] have showed, that the average missing transverse energy scale can be restored by calibrating jet energy scale in processes containing genuine $\mathrm{E_T^{miss}}$. The improvement to the $\mathrm{E_T^{miss}}$ scale from these Type-I $\mathrm{E_T^{miss}}$ corrections described in Section 5.3.1 is clearly visible in Fig. 6.20, which shows the ratio of the reconstructed $\mathrm{E_T^{miss}}$ to the true $\mathrm{E_T^{miss}}$ for the signal events with $\mathrm{m_A}$ = 500 GeV/c^2 after the requirement to have a third jet in the event as discussed in the previous sections. The true $\mathrm{E_T^{miss}}$ was built from stable Monte Carlo particles, excluding neutrinos and muons. The fitted mean value and standard deviation of the calibrated $\mathrm{E_T^{miss}}$ scale were found to be 1.01 and σ = 0.31, respectively. The tail on the righthand side of the $\mathrm{E_T^{miss}}$ ratio distributions was found to be caused by badly measured soft $\mathrm{E_T^{miss}}$ as can be seen from the dotted line which corresponds to the the calibrated $\mathrm{E_T^{miss}}$ with true $\mathrm{E_T^{miss}}$ less than 40 GeV.

To reconstruct the Higgs boson mass, the reconstructed energy of both τ leptons and τ neutrinos were required to be positive, i.e. $\mathrm{E^{reco}_{\tau_1,\tau_2}}$ >0 GeV and $\mathrm{E^{reco}_{\nu_1,\nu_2}}$ >0 GeV. The di-τ-lepton effective mass, $\mathrm{m_{\tau\tau}}$, can be calculated by applying the conservation of 4-momentum to the two-body decay of $\mathrm{H, A} \rightarrow \tau\tau$, which yields the expression

$$\mathrm{m^2_{\tau\tau}} = 2\mathrm{m^2_\tau} + 2\mathrm{E'_{\tau_1} E'_{\tau_2}} - 2\mathrm{p'_{\tau_1} p'_{\tau_2}} \cos\phi_{\tau_1,\tau_2}, \tag{6.1}$$

where ϕ_{τ_1,τ_2} is the angle between the two τ jet directions and where the prime denotes the τ lepton. Since $\mathrm{p_{\tau_{1,2}}} \gg \mathrm{m_\tau}$, $\mathrm{m_{\tau\tau}}$ can be expressed as

$$\mathrm{m_{\tau\tau}} = \sqrt{2(\mathrm{E_{\tau\ jet_1}} + \mathrm{E_{\nu_1}})(\mathrm{E_{\tau\ jet_2}} + \mathrm{E_{\nu_2}})(1 - \cos\Delta\phi)}, \tag{6.2}$$

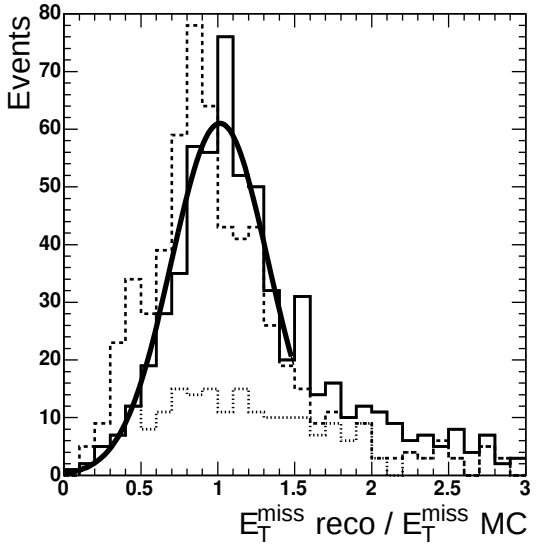

Figure 6.20: The calibrated (solid line) and non-calibrated (dashed line) ratio of the reconstructed $\mathrm{E_T^{miss}}$ and the true $\mathrm{E_T^{miss}}$ for the signal events with $\mathrm{m_A}$ = 500 GeV/c^2 after the requirement to have a third jet in addition to the two τ jets in the event. The dotted line corresponds to the ratio of the non-calibrated $\mathrm{E_T^{miss}}$, when the true $\mathrm{E_T^{miss}}$ is less than 40 GeV.

where $\mathrm{E}_{\tau\ \mathrm{jet}_{1,2}}$ are the reconstructed τ-jet energies, $\mathrm{E}_{\nu_{1,2}}$ are the reconstructed τ neutrino energies obtained with the collinearity approximation, and where $\Delta\phi$ is the total angle between the reconstructed τ jets.

The distribution of the reconstructed effective di-τ-lepton mass is shown in the left plot of Fig. 6.21 for the Higgs boson mass of $\mathrm{m_A}$ = 500 GeV/c^2 with calibrated $\mathrm{E_T^{miss}}$ and with different constraints. The mass peak is clearly visible, when the reconstructed energy of both neutrinos is required to be positive. It can be seen from equation (6.2), that the $\mathrm{m}_{\tau\tau}$ mass resolution depends on the resolution of $\mathrm{E_T^{miss}}$ and on the angle $\Delta\phi$ between the two τ jets in the transverse plane as $1/\sin(\Delta\phi)$. The larger the angle $\Delta\phi$ is, the worse the resolution becomes as shown in the right plot of Fig. 6.21 for events with positive neutrino energies. The large tail at the high $\mathrm{m}_{\tau\tau}$ values was found to be caused by the back-to-back di-τ-jet configuration, which is visible in the dashed histogram of the left plot of Fig. 6.21. The cut $\Delta\phi < 176^{\mathrm{o}}$ was found to partially remove the tail as can be seen in the solid histogram in the left plot of Fig. 6.21.

The distributions presented in Fig. 6.21 were produced without the requirement to have a third jet in the event. This requirement was observed to reduce much of the back-to-back di-τ-jet configurations, as demonstrated in Fig. 6.22 for signal events with $\mathrm{m_A}$ = 500 GeV/c^2.

The distribution of the reconstructed effective di-τ-lepton mass is shown for calibrated and non-calibrated $\mathrm{E_T^{miss}}$ in Fig. 6.23 for the signal with $\mathrm{m_A}$ = 500 GeV/c^2 for events with one additional jet. It can be seen, that the fraction of events with negative reconstructed neutrino energies was reduced, when the presence of the third jet was required. Furthermore, the tail reaching to high $\mathrm{m}_{\tau\tau}$ values was found to decrease with this requirement. When the left and right plots of Fig. 6.23 are compared, it is evident,

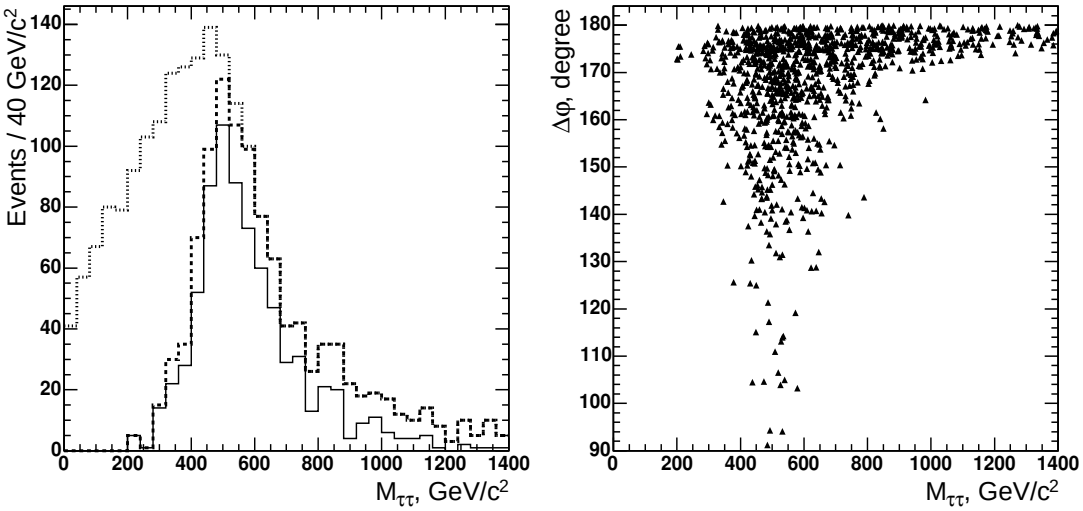

Figure 6.21: Left: the reconstructed effective di-τ-lepton mass with $\mathrm{E}^{\mathrm{reco}}_{\tau_1,\tau_2} > 0$ (dotted line), $\mathrm{E}^{\mathrm{reco}}_{\tau_1,\tau_2,\nu_1,\nu_2} > 0$ (dashed line), and with $\mathrm{E}^{\mathrm{reco}}_{\tau_1,\tau_2,\nu_1,\nu_2} > 0$ and $\Delta\phi$ <176° (solid line). Right: the scatter plot $\mathrm{m}_{\tau\tau}$ vs. $\Delta\phi$ for the Higgs boson of $\mathrm{m_A}$ = 500 GeV/c^2 with $\mathrm{E}^{\mathrm{reco}}_{\tau_1,\tau_2,\nu_1,\nu_2} > 0$. The plots have been produced without the requirement to have the third jet in the event.

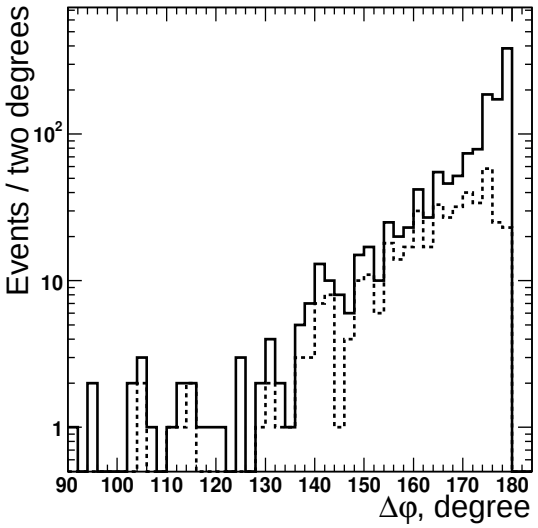

Figure 6.22: The angle $\Delta\phi$ between the τ jet directions before (solid line) and after (dashed line) the requirement to have a third jet in the event.

that applying the Type-I $\mathrm{E}^{\mathrm{miss}}_{\mathrm{T}}$ corrections improves significantly the Higgs boson mass reconstruction efficiency, the mean value of the Higgs boson mass distribution and the mass resolution. The Higgs boson mass reconstruction efficiency, the fitted mean value of the $\mathrm{m}_{\tau\tau}$ distribution and the mass resolution are summarized in Table 6.4 for the signal with $\mathrm{m_A}$ = 500 GeV/c^2. The $\mathrm{m}_{\tau\tau}$ mass resolution obtained with the calibrated $\mathrm{E}^{\mathrm{miss}}_{\mathrm{T}}$ was found to be 17.6 (18.6)% for the signal with $\mathrm{m_A}$ = 200 (800) GeV/c^2, respectively.

The reconstructed $\mathrm{m}_{\tau\tau}$ distributions for the QCD multi-jet events generated within the different $\hat{\mathrm{p}}_{\mathrm{T}}$ bins

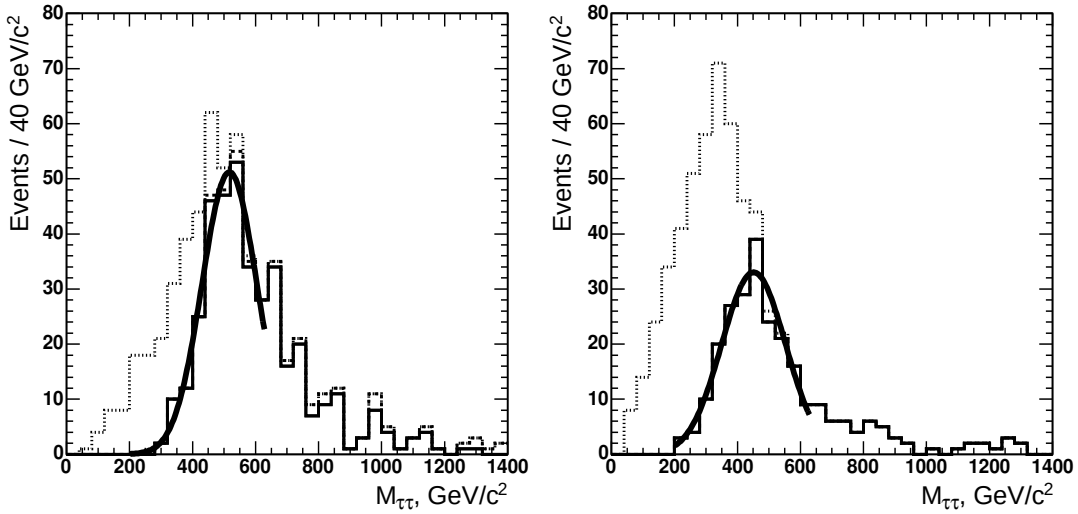

Figure 6.23: The reconstructed effective di-τ-lepton mass distributions with calibrated (left plot) and non-calibrated (right plot) $\mathrm{E_T^{miss}}$ for the signal with $\mathrm{m_A}$ = 500 GeV/c^2 and a third jet in the event. The dotted, dashed, and solid histograms corresponds to the $\mathrm{E^{reco}_{\tau_1,\tau_2}} > 0$, $\mathrm{E^{reco}_{\tau_1,\tau_2,\nu_1,\nu_2}} > 0$, and $\mathrm{E^{reco}_{\tau_1,\tau_2,\nu_1,\nu_2}} > 0$ with $\Delta\phi$ <176° requirements, respectively.

	Efficiency of $\mathrm{m}_{\tau\tau}$ reconstruction, $\mathrm{E^{reco}_{\tau_1,\tau_2}} > 0$	$\mathrm{E^{reco}_{\nu_1,\nu_2}} > 0$	$\langle \mathrm{m^{reco}_{\tau\tau}} \rangle$, GeV/$c^2$	$\sigma(\mathrm{m^{reco}_{\tau\tau}})$, GeV/$c^2$	$\sigma(\mathrm{m^{reco}_{\tau\tau}})/\mathrm{m^{sim}_{\tau\tau}}$, %
Non-calibrated $\mathrm{E_T^{miss}}$	0.93	0.43	470	97	19.4
Calibrated $\mathrm{E_T^{miss}}$	0.93	0.67	520	89	17.8

Table 6.4: The performance of the Higgs boson mass reconstruction with uncalibrated and Type-I calibrated $\mathrm{E_T^{miss}}$. The columns two and three show the efficiencies of Higgs boson mass reconstruction; the fourth and fifth columns show the fitted mean and values of the $\mathrm{m}_{\tau\tau}$ mass distribution and the final column shows the $\mathrm{m}_{\tau\tau}$ resolution. The results are presented for the Higgs boson with $\mathrm{m_A}$ = 500 GeV/c^2.

are shown in Fig. 6.24. The events were required to contain the third jet in addition to the two τ-jet candidates. Since the difference between requiring positive neutrino energies and requiring positive neutrino energies with $\Delta\phi$ <176° was observed to be small in the Figs. 6.23 and 6.24, it was decided to use only the $\mathrm{E^{reco}_{\nu_1,\nu_2}} > 0$ condition when reconstructing the Higgs boson mass. The reconstructed Higgs boson mass values were required to be within the mass windows of 150-300 GeV/c^2, 400-700 GeV/c^2, and 600-1100 GeV/c^2 for the search for the signal of $\mathrm{m_A}$ = 200, 500, and 800 GeV/c^2, respectively.

6.5 Results

In the following, the efficiency of the event selections described in the Sections 6.3 and 6.4, the corresponding cross-sections and the expected number of events after the event selections are summarized for the signal and the background processes. The efficiency of each selection shown in the summary tables

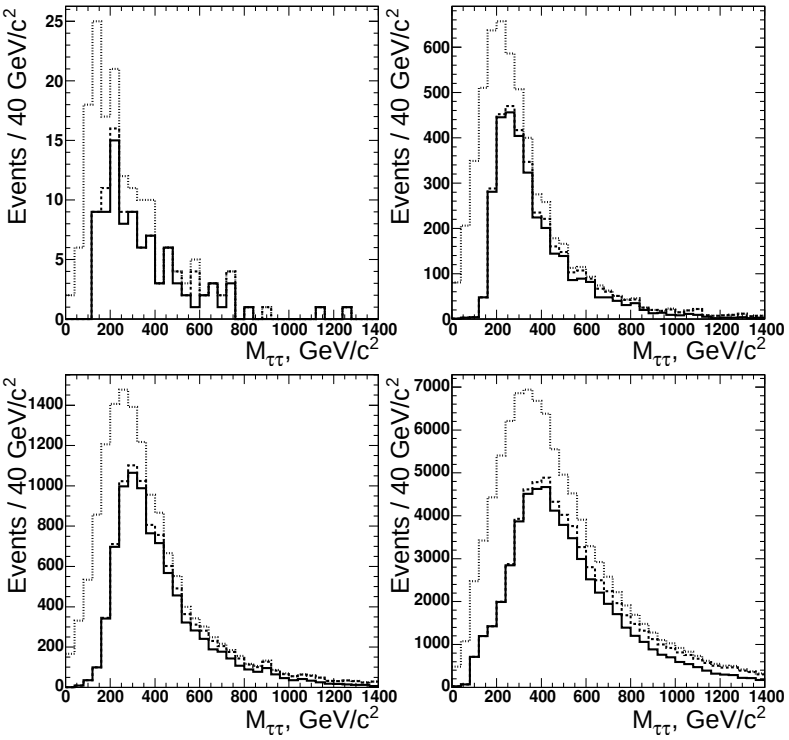

Figure 6.24: The reconstructed effective di-τ-lepton mass distributions with calibrated $\mathrm{E_T^{miss}}$ and with the third jet required in the event for QCD multi-jet events with $50 < \hat{\mathrm{p}}_\mathrm{T} < 80$ GeV/c (top left), $80 < \hat{\mathrm{p}}_\mathrm{T} < 120$ GeV/c (top right), $120 < \hat{\mathrm{p}}_\mathrm{T} < 170$ GeV/c (bottom left), and $\hat{\mathrm{p}}_\mathrm{T} > 170$ GeV/c (bottom right). The dotted, dashed, and solid histograms corresponds to the $\mathrm{E}^{\mathrm{reco}}_{\tau_1,\tau_2} > 0$, $\mathrm{E}^{\mathrm{reco}}_{\tau_1,\tau_2,\nu_1,\nu_2} > 0$, and $\mathrm{E}^{\mathrm{reco}}_{\tau_1,\tau_2,\nu_1,\nu_2} > 0$ with $\Delta\phi$ <176$^\mathrm{o}$ requirements, respectively.

was evaluated relative to the previous selection.

The expectations for the signal of $\mathrm{m_A}$ = 200, 500, and 800 GeV/c^2 are summarized in Table 6.5. The signal cross-sections and the branching ratios were evaluated for the $\mathrm{m_h^{max}}$-scenario with μ = 200 GeV/c^2. It should be noted, that for $\mathrm{m_A}$ = 200 GeV/c^2 both one and three tracks were allowed for the number of tracks inside the signal cone, where as only one track was allowed inside the signal cone for $\mathrm{m_A}$ = 500 and 800 GeV/c^2. The number of events expected for an integrated luminosity of 60 fb^{-1} is 58.0, 27.0, and 11.0 for $\mathrm{m_A}$ values of 200 ($\tan\beta$ = 20), 500 ($\tan\beta$ = 30), and 800 GeV/c^2 ($\tan\beta$ = 40), respectively.

Despite the more than one million generated events and the generation preselection, the number of QCD multi-jet events was not enough to ensure a large number of Monte Carlo events passing all the event selections. In order to decrease the statistical uncertainties, the selection criteria were factorized into three groups for evaluating the QCD multi-jet background.

The first group (group 1) included the Level-1 trigger and the calorimetric reconstruction of the jets (at HLT and offline). It included also the cut on the transverse energy of the jets. After the event passed the group 1 selection, the two other selection groups (group 2 and group 3) were applied independently. The second group contained essentially the τ-jet identification part of the analysis, i.e. the charged track isolation (at both HLT and offline), the cut on the $\mathrm{p_T}$ of the leading track and the selection on the number

	$\mathrm{m_A}$ = 200 GeV/c^2 $\tan\beta$ = 20	$\mathrm{m_A}$ = 500 GeV/c^2 $\tan\beta$ = 30	$\mathrm{m_A}$ = 800 GeV/c^2 $\tan\beta$ = 40
Cross-sections and branching ratios			
$\sigma(\mathrm{gg} \rightarrow \mathrm{b\bar{b}(A+H)})$ (fb)	45795+44888	2741+2744	677+677
$\mathrm{BR(H, A} \rightarrow \tau\tau)$	0.1	0.082	0.087
$\mathrm{BR}(\tau \rightarrow \mathrm{hadr.})$	0.65×0.65	0.65×0.65	0.65×0.65
$\sigma \times \mathrm{BR}$ (fb)	3831	190	49.8
Experimental selection efficiencies			
Level-1 trigger	0.506	0.854	0.896
High-level trigger	0.289	0.319	0.314
Two offline calorimeter τ jets	0.997	0.999	0.999
Cuts on $\mathrm{E_T}$ of τ jets	0.430	0.755	0.780
Two offline τ-jet candidates	0.674	0.716	0.675
$\mathrm{p_T^{ltr}} > 35$ GeV/c	0.326	0.616	0.713
Charged track isolation	0.859	0.950	0.954
$\mathrm{N_{tracks}}$ in signal cone	0.81	0.67	0.78
$\mathrm{Q_1 \times Q_2 = -1}$	0.98	0.94	0.94
$\geq$1 extra jet, $\mathrm{E_T^{raw}} > 20$ GeV	0.21	0.27	0.31
Only 1 extra jet, $\mathrm{E_T^{raw}} > 20$ GeV	0.83	0.82	0.78
b tagging of the extra jet	0.36	0.44	0.41
$\mathrm{m_{\tau\tau}}$ reconstruction efficiency			
$\mathrm{E^{reco}_{\tau_1,\tau_2}} > 0$	0.93	0.93	0.92
$\mathrm{E^{reco}_{\nu_1,\nu_2}} > 0$	0.56	0.67	0.67
$\mathrm{m_{\tau\tau}}$ mass window	150-300 GeV/c^2	400-700 GeV/c^2	600-1100 GeV/c^2
Mass window efficiency	0.81	0.73	0.81
Summary of selection			
Total efficiency	2.5×10^{-4}	2.4×10^{-3}	3.6×10^{-3}
Cross-section after selections (fb)	0.96	0.46	0.19
Number of events at 60 fb^{-1}	58.0	27.0	11.0

Table 6.5: The summary table of the selections for the signal of $\mathrm{m_A}$=200, 500, and 800 GeV/c^2.

of tracks inside the signal cone. The third group included the selection of the third (labeled as extra) jet in the event, the tagging of the third jet as a b jet, and the di-τ-jet mass reconstruction. The choice of the selections in the second and third groups was made by minimizing the correlation between the groups. Furthermore, the selection criteria for the number of tracks were factorized inside group 2 and the requirement of positive neutrino energies and the selection criteria on the $\mathrm{m_{\tau\tau}}$ reconstruction and b-jet tagging were factorized inside group 3.

The results for the QCD multi-jet background are summarized in Table 6.6 for the signal selection with $\mathrm{m_A}$ = 200 GeV/c^2. The requirement, that the τ jets have opposite total charges ($\mathrm{Q_1 \times Q_2 = -1}$), was not included in the table. It was observed to reduce the QCD multi-jet background by another factor of two, leading to 104 events for the QCD multi-jet background expected for an integrated luminosity of 60 fb^{-1}.

The difference of the analysis for the $\mathrm{m_A}$ = 200 GeV/c^2 and the heavier $\mathrm{m_A}$ values was, that for the heavier $\mathrm{m_A}$ values only one track was allowed in the signal cone instead of one or three. Furthermore, the jet energy thresholds and Higgs boson mass windows were set separately for each $\mathrm{m_A}$ value. When the selections were applied to search for the signal with $\mathrm{m_A}$ = 500 GeV/c^2 and $\mathrm{m_A}$ = 800 GeV/c^2, the expected number of the QCD multi-jet events was found to be 25.0 and 4.0, respectively, for an integrated luminosity of 60 fb^{-1}.

	QCD di-jet background in bins of generated $\hat{\mathrm{p}}_\mathrm{T}$			
	$>$170 GeV/c	120-170 GeV/c	80-120 GeV/c	50-80 GeV/c
Cross-sections and kinematical preselections				
σ (fb)	1.33×10^{8}	5.03×10^{8}	2.94×10^{9}	2.08×10^{10}
$\epsilon_{\mathrm{kine\ presel.}}$	2.12×10^{-1}	4.19×10^{-2}	5.77×10^{-3}	2.44×10^{-4}
Group 1 cuts: L1 trigger + L2 and offline calo reco + $\mathrm{E_T}$cut				
Level-1 trigger	0.562	0.726	0.715	0.461
Two HLT jets, $\Delta\mathrm{R_{j_1,j_2}} > 1.0$	0.927	0.959	0.982	0.987
Two offline calorimeter τ jets	0.975	0.975	0.982	0.994
Cuts on $\mathrm{E_T}$ of τ jets	0.753	0.804	0.774	0.343
$\epsilon_{\mathrm{group1}}$	0.383	0.547	0.534	0.155
Group 2 cuts: τ-jet identification at HLT and offline				
HLT calo + Pxl τ trigger	7.15×10^{-4}	1.81×10^{-3}	4.44×10^{-3}	1.12×10^{-2}
Two offline τ-jet candidates	0.86	0.84	0.825	0.84
$\mathrm{p_T^{ltr}} > 35$ eV/c	0.47	0.41	0.42	0.38
Charged track isolation	0.24	0.21	0.25	0.35
1 or 3 prongs in first τ jet	0.66	0.92	0.63	0.72
1 or 3 prongs in second τ jet	0.48	0.54	0.65	0.72
$\epsilon_{\mathrm{group2}}$ after passing group1	2.30×10^{-5}	6.33×10^{-5}	1.63×10^{-4}	6.54×10^{-4}
Group 3 cuts: extra jet reconstruction, b tagging and mass window				
$\geq$1 extra jet, $\mathrm{E_T^{raw}}$>20 GeV	0.463	0.235	0.127	0.090
Only 1 extra jet, $\mathrm{E_T^{raw}}$>20 GeV	0.661	0.817	0.863	0.855
$\mathrm{E^{reco}_{\tau_1,\tau_2}} > 0$	0.921	0.898	0.882	0.834
$\mathrm{E^{reco}_{\nu_1,\nu_2}} > 0$	0.701	0.683	0.657	0.625
b tagging of the extra jet	0.098	0.050	0.033	0.016
$\mathrm{m_{\tau\tau}}$window: 150-300 GeV/c^2	0.142	0.295	0.433	0.430
$\epsilon_{\mathrm{group3}}$ after passing group1	2.77×10^{-3}	1.75×10^{-3}	9.15×10^{-4}	2.28×10^{-4}
Summary of selection				
$\epsilon_{\mathrm{group1}} \times \epsilon_{\mathrm{group2}} \times \epsilon_{\mathrm{group3}}$	2.44×10^{-8}	6.07×10^{-8}	7.98×10^{-8}	2.84×10^{-8}
Cross-section after selections (fb)	0.69	1.28	1.35	0.144
Number of events at 60 fb^{-1}	41.4	76.7	81.2	8.7

Table 6.6: The summary table of the selections for the QCD multi-jet background for the search of Higgs bosons with $\mathrm{m_A}$ = 200 GeV/c^2. The selections are factorized as explained in the text. The requirement of the τ-jet candidates to have opposite charge ($\mathrm{Q_{\tau 1} \times Q_{\tau 2} = -1}$) has not been included in the table.

The total number of events in the electroweak backgrounds passing all selections was found to be small compared to the QCD multi-jet background. The expected number of events from the irreducible

(i.e. non-suppressed) electroweak backgrounds is shown in Table 6.7 for the search for the signal of $\mathrm{m_A}$ = 200 GeV/c^2 and for an integrated luminosity of 60 fb^{-1}. The efficiencies for some of the event selections are also shown in the table.

In total, 6.0 events are expected for electroweak backgrounds for the search of Higgs bosons with $\mathrm{m_A}$ = 200 GeV/c^2 and for an integrated luminosity of 60 fb^{-1}. The number of events in the irreducible electroweak backgrounds passing all event selections for the signal of $\mathrm{m_A}$ = 500 and 800 GeV/c^2 was found to be 4.0 and 1.0, respectively, for an integrated luminosity of 60 fb^{-1}.

Process	$\mathrm{N_{exp.}}$ at 60 fb^{-1}	$\mathrm{Q_{\tau 1} \times Q_{\tau 2}}$ = -1	Only one extra jet	b tagged jet	$\mathrm{m_{\tau\tau}}$ window
$\mathrm{t\bar{t}}$	0.64	0.96	0.36	0.42	0.11
W+jet	0.33	0.81	0.15	0.06	0.12
Wt	0.26	0.96	0.49	0.44	0.23
$\mathrm{Z}/\gamma^* \to \tau\tau$, $130 < \mathrm{m_{\tau\tau}} < 300$ GeV/c^2	3.80	0.96	0.23	0.06	0.61
$\mathrm{Z}/\gamma^* \to \tau\tau$, $\mathrm{m_{\tau\tau}} > 300$ GeV/c^2	0.18	0.95	0.27	0.05	0.04
$\tau\tau\mathrm{b\bar{b}}$, $\mathrm{m_{\tau\tau}} > 100$ GeV/c^2	0.86	0.98	0.39	0.44	0.38

Table 6.7: The number of expected events for an integrated luminosity of 60 fb^{-1} and efficiencies of some of the selections for the reducible backgrounds.

The expected $\mathrm{m_{\tau\tau}}$ distributions for signal with $\mathrm{m_A}$ = 200 GeV/c^2 ($\tan\beta$ = 20) and $\mathrm{m_A}$ = 500 GeV/c^2 ($\tan\beta$ = 30) are shown in Fig. 6.25 for an integrated luminosity of 60 fb^{-1} in the $\mathrm{m_h^{max}}$ MSSM scenario [39, 127].

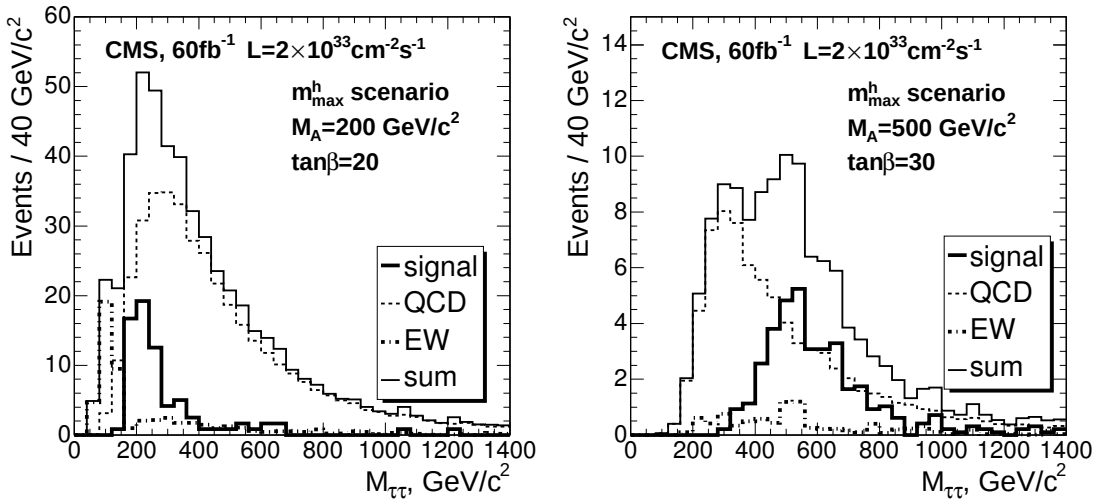

Figure 6.25: The expected $\mathrm{m_{\tau\tau}}$ distributions for the signal of $\mathrm{m_A}$ = 200 GeV/c^2, $\tan\beta$ = 20 (left plot) and $\mathrm{m_A}$ = 500 GeV/c^2, $\tan\beta$ = 30 (right plot), and the backgrounds for an integrated luminosity of 60 fb^{-1}. The thick solid, dashed, and dashed-dotted lines correspond to the signal, QCD multi-jet background, and electroweak background events. The sum of the signal and background events is visualized with the thin solid line.

6.6 Systematic uncertainties and evaluation of the background from data

In the following subsections, the systematic uncertainties affecting the neutral Higgs boson mass distribution are discussed. Furthermore, a method is presented for the evaluation of the QCD multi-jet background from the data. The systematic uncertainties resulting from the tracker misalignment and the plans for measuring the jet $\rightarrow \tau$ fake rate have been discussed earlier in Section 5.1.13.

6.6.1 Effect of $\mathrm{E_T^{miss}}$ and jet energy scale uncertainties

The effect of the $\mathrm{E_T^{miss}}$ and the jet energy scale uncertainty on the neutral Higgs boson mass reconstruction efficiency was estimated with the Type-I $\mathrm{E_T^{miss}}$ corrections described in Section 5.3.1. For the low $\mathrm{E_T}$ part ($\mathrm{E_T} < 50$ GeV), the scale uncertainty of 10% was applied, while a 3% uncertainty was used for the high $\mathrm{E_T}$ part ($\mathrm{E_T} > 50$ GeV). The variation of the scale was applied independently for the two parts to obtain the maximal deviations from the case with no uncertainty.

The efficiency of the neutral Higgs boson mass reconstruction and the neutral Higgs boson mass measurement is summarized in Table 6.8 taking into account the $\mathrm{E_T^{miss}}$ and the jet energy scale uncertainties for the signal with $\mathrm{m_A}$ = 500 GeV/c^2. The efficiency for the $\mathrm{E^{reco}_{\nu_1,\nu_2}} > 0$ selection was found to deviate in the worst case by $\sim$3% from the nominal efficiency. The deviation was caused in this case by the combination of $\pm$3% shift of the jet scale and -10% shift of the $\mathrm{E_T^{miss}}$ scale. The mean value of the $\mathrm{m_{\tau\tau}}$ fit was found to vary within +16 GeV/c^2 and -10 GeV/c^2 from the nominal value of 520 GeV/c^2. The largest contribution was observed to be caused by the $\mathrm{E_T^{miss}}$ scale uncertainty.

Jet scale	$\mathrm{E_T^{miss}}$ scale	$\mathrm{E^{reco}_{\tau_1,\tau_2}} > 0$	$\mathrm{E^{reco}_{\nu_1,\nu_2}} > 0$	$\langle \mathrm{m_{\tau\tau}{}^{reco}} \rangle$, GeV/$c^2$
No uncertainty		0.93$\pm$0.01	0.67$\pm$0.02	520$\pm$8
+3%	0%	0.93$\pm$0.01	0.67$\pm$0.02	536$\pm$10
-3%	0%	0.93$\pm$0.01	0.67$\pm$0.02	518$\pm$9
+3%	+10%	0.93$\pm$0.01	0.67$\pm$0.02	532$\pm$9
+3%	-10%	0.93$\pm$0.01	0.65$\pm$0.02	530$\pm$9
-3%	+10%	0.93$\pm$0.01	0.67$\pm$0.02	521$\pm$9
-3%	-10%	0.93$\pm$0.01	0.65$\pm$0.02	510$\pm$8

Table 6.8: The efficiency of the neutral Higgs boson mass reconstruction and the neutral Higgs boson mass measurement taking into account the $\mathrm{E_T^{miss}}$ and the jet energy scale uncertainties the for the signal with $\mathrm{m_A}$ = 500 GeV/c^2.

6.6.2 Measurement of the QCD multi-jet background from the data

The QCD multi-jet production was found to be the largest background for the $\mathrm{H, A} \rightarrow \tau\tau$ channel. The following method is proposed to evaluate this background from the data.

A control sample must be used where all signal selections are applied except for the mass window cut. Instead of requiring the τ-jet candidates to have opposite charge, the sample is proposed to be selected by requiring the τ-jet candidates to have the same sign (same sign sample). Since the contamination of the signal events and irreducible backgrounds is negligible in the same sign sample, this approach yields the opportunity to predict the QCD multi-jet background from the data in a given mass window deducted from the number of events and the measured shape of di-τ-jet mass in the same sign sample. The expected number of the same signed QCD multi-jet events after all selections except for the mass window for the signal with $\mathrm{m_A}$ = 200 GeV/c^2 is 380 for an integrated luminosity of 60 fb^{-1}. Neglecting the uncertainty of the measured shape of the di-τ-jet mass, this approach leads to a 5% statistical uncertainty of the QCD multi-jet background estimates in the signal mass window. For the $\mathrm{m_A}$ = 500 (800) GeV/c^2 selections about 80 (28) of the same signed QCD multi-jet events are expected, which yields a statistical uncertainty of $\sim$10 (20)%.

6.7 Discovery reach in the $\mathrm{m_A}$,$\tan\beta$-plane

The lowest value of $\tan\beta$ for which the 5σ discovery is possible with an integrated luminosity of 60 fb^{-1} is shown in Table 6.9 for the three $\mathrm{m_A}$ values considered in the $\mathrm{H, A} \rightarrow \tau\tau$ analysis. The values are shown with and without the systematic uncertainty of the QCD multi-jet background. The 5σ discovery contour is shown in Fig. 6.26 for an integrated luminosity of 60 fb^{-1} without the systematic uncertainties. The significance of the discovery was calculated with Poissonian statistics with the $\mathrm{S_{cP}}$ program [129].

The extension of the discovery reach to lower $\tan\beta$ values would be possible with a lower threshold on the transverse energy of the third jet in the event, provided that the fake jets would then be suppressed with jet to track matching. Another improvement is expected from the increase of the neutral Higgs boson mass reconstruction efficiency with the improved missing $\mathrm{E_T^{miss}}$ measurement from the particle flow algorithm described in Section 5.1.14.1. Furthermore, advances in the b-jet tagging performance are expected to further extend the discovery reach to the lower the $\tan\beta$ threshold.

Low $\tan\beta$ limit for 5σ discovery	Higgs boson mass $\mathrm{m_A}$ = 200 GeV/c^2	$\mathrm{m_A}$ = 500 GeV/c^2	$\mathrm{m_A}$ = 800 GeV/c^2
No systematics	20	32	46
With systematics	21	34	49

Table 6.9: The low limit of $\tan\beta$ for which the 5σ discovery is possible with data of 60 fb^{-1}.

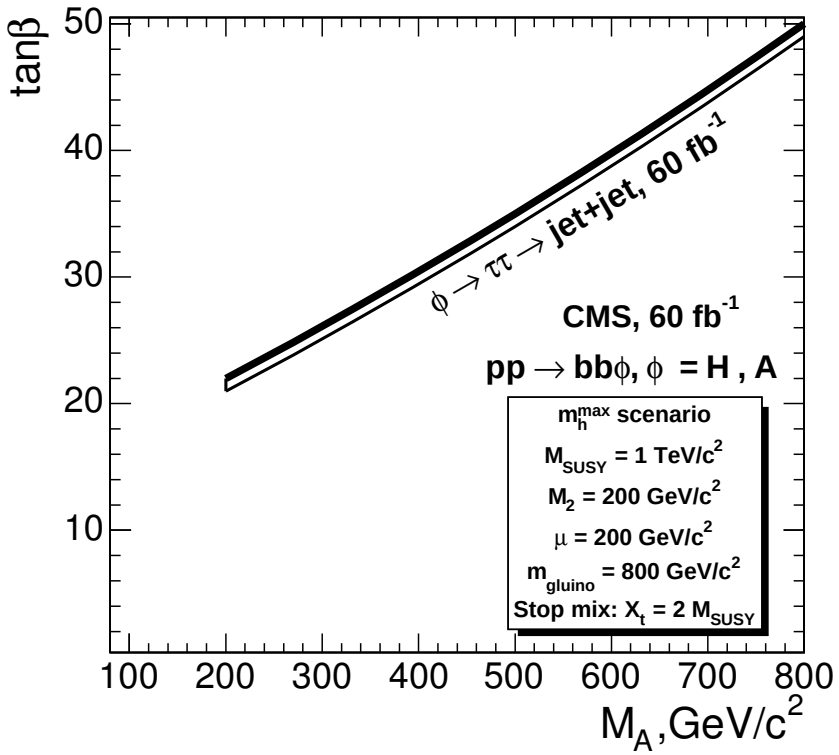

Figure 6.26: The 5σ discovery contours for the $\mathrm{gg} \rightarrow \mathrm{b\bar{b}H(A)}$, $\mathrm{H, A} \rightarrow \tau\tau \rightarrow \mathrm{hadr.}$ decay channel in the $\mathrm{m_h^{max}}$ scenario for an integrated luminosity of 60 fb^{-1} [25].

6.8 Further developments after the publication of the physics TDRs

After the publication of the physics TDRs, the attention of the analyses with τ identification has been shifted to LHC startup scenarios with the expected LHC luminosity of 2×10^{30} cm^{-2}s^{-1} and 10^{31} cm^{-2}s^{-1}. Since the discovery of the neutral Higgs bosons would require an integrated luminosity of tens of fb^{-1} in the $\mathrm{gg} \rightarrow \mathrm{b\bar{b}H(A)}$, $\mathrm{H, A} \rightarrow \tau\tau \rightarrow \mathrm{hadr.}$ process, no new study has been made of this process after the publication of the physics TDRs.

7 Search for heavy charged Higgs bosons in the $\mathrm{H}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau} \rightarrow \mathrm{hadr.}$ channel

As described in Section 2.2.5, the charged MSSM Higgs bosons are produced at the LHC for $\mathrm{m}_{\mathrm{H}^{\pm}} > \mathrm{m}_{\mathrm{t}}$ mainly through the $\mathrm{gb} \rightarrow \mathrm{tH}^{\pm}$ and $\mathrm{gg} \rightarrow \mathrm{tbH}^{\pm}$ processes. The main decay channel $\mathrm{H}^{\pm} \rightarrow \mathrm{tb}$ has been shown to be difficult to observe [25]. The subleading decay mode $\mathrm{H}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$, discussed in this chapter, is interesting and usable because of the tight τ-jet identification, the possibility of exploiting the τ helicity correlations, and the possibility of reconstructing the transverse Higgs boson mass in fully hadronic events. After a veto on isolated leptons, identification of the hadronic τ decay is applied with the methods described in Section 5.1. Additionally, the missing E_{T} measurement, b-tagging of one of the associated b jets, the reconstruction of the associated top mass, and a mass window on the $\mathrm{m}_{\tau\nu_{\tau}}$ transverse mass are applied. The spin-parity correlations of the $\mathrm{H}^{\pm}$ and $\mathrm{W}^{\pm}$ give separation power to distinguish signal τ jets from τ jets coming from the $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decays in the electroweak backgrounds as described in Section 2.3.2.

In the following, the event samples used for the analysis are described in Section 7.1, the trigger and offline event selections are presented in Section 7.2 and 7.3, respectively, and the results are shown in Section 7.4. The systematic effects affecting the results and methods used to measure the background from data are discussed in Section 7.5. The discovery reach in the m_{A},$\tan\beta$ plane is presented in Section 7.6. Finally, recent developments after the publication of the physics TDRs are described in Section 7.7.

The methods and results presented in Sections 7.1-7.6 are based on Refs. [25, 130]. After these studies, the τ identification part was revisited and repeated with improved optimization. The results of the reoptimized τ identification are presented in Section 7.7.1.

7.1 Simulation of event samples

The signal samples were generated with the PYTHIA [69] Monte Carlo generator process $\mathrm{gg} \rightarrow \mathrm{tbH}^{\pm}$ in the $\mathrm{m}_{\mathrm{h}}^{\mathrm{max}}$ MSSM scenario [39, 127] with the following parameters: M_2 = 200 GeV/c^2, μ = 200 GeV/c^2, M_3 = 800 GeV/c^2, $\mathrm{M}_{\mathrm{SUSY}}$ = 1 TeV/c^2, X_{t} = 2 TeV/c^2, m_{t} = 175 GeV/c^2, and $\tan\beta$ = 30. As described in Section 2.2.5, the processes $\mathrm{gb} \rightarrow \mathrm{tH}^{\pm}$ and $\mathrm{gg} \rightarrow \mathrm{tbH}^{\pm}$ contribute to the charged Higgs boson production at $\mathrm{m}_{\mathrm{H}^{\pm}} > \mathrm{m}_{\mathrm{t}}$. The production cross-section of the former of these processes is by a factor $\sim$2-3 larger than the latter one. It was, however, not possible to add the production mode $\mathrm{gb} \rightarrow \mathrm{tH}^{\pm}$ to the $\mathrm{gg} \rightarrow \mathrm{tbH}^{\pm}$ process, since it would have required weighting of the events in the full event simulation, which was not available in the software used. Near the top mass threshold, $\mathrm{m}_{\mathrm{H}^{\pm}} \sim \mathrm{m}_{\mathrm{t}}$, only the exclusive process $\mathrm{gg} \rightarrow \mathrm{tbH}^{\pm}$ can lead to a correct event description. Therefore, only the $\mathrm{gg} \rightarrow \mathrm{tbH}^{\pm}$ process was used for the event generation. The signal cross-sections were normalized to the NLO cross-sections

calculated from Ref. [131] for the different $\mathrm{m_A/m_{H^{\pm}}}$ values. The TAUOLA package [74] was used to handle the τ decays in order to include the τ polarization to the decays. The τ's from the $\mathrm{H}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ process were forced to decay hadronically.

The most dangerous background in terms of the cross-section is the QCD multi-jet background which may fake both τ jets and $\mathrm{E_T^{miss}}$. The QCD multi-jet samples were generated with PYTHIA with $120 < \hat{\mathrm{p}}_{\mathrm{T}} < 380$ GeV/c. Other dangerous backgrounds are the $\mathrm{t\bar{t}}$, Wt, and W+3 jet events. These backgrounds contain genuine τ jets and $\mathrm{E_T^{miss}}$ which might be confused as signal. They also contain hadronic jets, which may fake τ jets and/or $\mathrm{E_T^{miss}}$ as in the QCD multi-jet events. In addition, the second top quark of the $\mathrm{t\bar{t}}$ events may produce neutrinos through leptonic decays of the W boson contaminating thus the $\mathrm{E_T^{miss}}$ measurement. The $\mathrm{t\bar{t}}$, Wt, and W+3 jet backgrounds were generated with Pythia, TopRex [76], and MadGraph [73], respectively. The $\mathrm{t\bar{t}}$ cross-section was normalized to the NLO cross-section. The TAUOLA package was used to simulate the τ decays. All τ decays were allowed for the backgrounds.

Because of the large number of expected background events, preselections were applied at the generator level for the $\mathrm{t\bar{t}}$ and Wt backgrounds. These backgrounds were required to contain at least one jet with $\mathrm{E_T} > 80$ GeV in a cone of 0.5 and at least one charged hadron with $\mathrm{p_T} > 60$ GeV/c. The preselection efficiency for the $\mathrm{t\bar{t}}$ and Wt backgrounds was 5.2% and 2.5%, respectively.

The response of the CMS detector was simulated with the CMSIM [79] and OSCAR [78] packages with pile-up events corresponding to the luminosity of $\mathcal{L} = 2\times10^{33}$ cm^{-2}s^{-1}. The events were reconstructed with the ORCA [80] software version 8.7.4. The cross-sections and the products of the cross-section and branching ratio of the signal and background samples are summarized in Table 7.1.

Process	Generator	$\mathrm{m_A}$ (GeV/c^2)	$\mathrm{m_{H^{\pm}}}$ (GeV/c^2)	σ (fb)	$\sigma \times \mathrm{BR}$ (fb)
$\mathrm{gg} \rightarrow \mathrm{tbH}^{\pm}$, $\mathrm{H}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$	PYTHIA	183	201.0	1589	775.5
$\mathrm{gg} \rightarrow \mathrm{tbH}^{\pm}$, $\mathrm{H}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$	PYTHIA	289	300.9	687	118.3
$\mathrm{gg} \rightarrow \mathrm{tbH}^{\pm}$, $\mathrm{H}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$	PYTHIA	392	400.7	320	37.7
$\mathrm{t\bar{t}}$	PYTHIA			8.40×10^5	1.24×10^5
Wt	TopRex			6.2×10^4	9.14×10^3
W^++3 jets	MadGraph			2.09×10^6	2.47×10^5
W^-+3 jets	MadGraph			1.46×10^6	1.72×10^5
QCD, $\hat{\mathrm{p}}_{\mathrm{T}}$=120-380 GeV/$c$	PYTHIA			6.4×10^8	6.4×10^8

Table 7.1: Cross-section and branching fraction times cross-section for the signal and background samples with $\tan\beta = 30$.

7.2 Trigger level selections

Since the heavy charged Higgs bosons with $\mathrm{m_{H^{\pm}}} > \mathrm{m_t}$ and $\mathrm{H}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau} \rightarrow \mathrm{hadr.}$ produce a hard τ jet, the $\mathrm{H}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ events can be triggered with the single τ trigger. The Level-1 τ-jet candidate described in Section 5.1.3 was required to exceed 93 GeV. After the Level-1 trigger, the Trk+Tau high level trigger with charged track isolation based on regional track reconstruction was applied as described in Section 5.1.4.2. Tracks consisting of maximally six hits were reconstructed around the Level-1 τ-jet axis before applying charged track isolation described in Section 5.1.6. The transverse momentum of the leading track was required to exceed 25 GeV/c to suppress QCD multi-jet events. The charged track isolation was applied with the parameters $\mathrm{R_m}$ = 0.10, $\mathrm{R_s}$ = 0.065, $\mathrm{R_i}$ = 0.40, and $\mathrm{p_T^{min}}$ = 1.0 GeV/c. In addition to the charged tracker isolation, a cut was placed on the uncorrected missing $\mathrm{E_T}$. The uncorrected missing $\mathrm{E_T}$ was required to exceed 67 GeV for the signal and electroweak backgrounds and 60 GeV for the QCD multi-jet background.

The trigger efficiencies and the contamination from electrons are shown in Tables 7.2, 7.3 and 7.4 for the signal and the electroweak and QCD multi-jet backgrounds, respectively. For the signal events, more than 80% of the HLT τ jets were found to be true τ's coming from the $\mathrm{H}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decay. Since the signal τ's were forced to decay hadronically, less than 1% of signal events passing the single τ trigger were observed to contain a leading electron in the τ jet as a result of photon conversions from the $\pi^0 \rightarrow \gamma\gamma$ decays. For the electroweak backgrounds, in which the τ decays were not forced, a considerable fraction of the events passing the single τ trigger were found to contain a $\tau \rightarrow \mathrm{e}\nu_{\mathrm{e}}\nu_{\tau}$ decay.

	$\mathrm{m_{H^{\pm}}}$ (GeV/c^2)		
	200	300	400
L1 trigger ($\mathrm{E_T} > 93$ GeV)	58.2%	69.9%	75.6%
$\mathrm{E_T^{miss}} > 67$ GeV	44.9%	62.4%	74.7%
Isolation, $\mathrm{p_T} > 25$ GeV/c	47.3%	55.8%	56.7%
Total efficiency	12.3%	24.3%	32.0%
τ jet purity	87.1%	92.4%	94.5%
Electron contamination	0.6%	0.7%	0.9%

Table 7.2: Efficiencies of the Level-1 and high level triggers, fraction of MC matched τ jets of all τ jet candidates (τ-jet purity) and electron contamination for the signal events with $\mathrm{m_{H^{\pm}}}$ = 200-400 GeV/c^2 and $\tan\beta$ = 30.

7.3 Offline event reconstruction and selections

After applying the Level-1 and HLT triggers to the events, the events were required to pass a series of offline event selections. To ensure, that the missing energy in the event is coming from the τ jet from the $\mathrm{H}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decay, a lepton veto was applied to remove events which contained isolated electrons

	$\mathrm{t\bar{t}}$	Wt	W+3 jets
Preselection	5.2%	2.6%	-
L1 trigger ($\mathrm{E_T} > 93$ GeV)	73.4%	78.1%	29.8%
$\mathrm{E_T^{miss}} > 67$ GeV	41.8%	23.1%	55.0%
Isolation, $\mathrm{p_T} > 25$ GeV/c	16.2%	28.9%	14.7%
Total efficiency	2.59×10^{-3}	2.27×10^{-3}	1.0%
Electron contamination	18.9%	16.4%	30.5%

Table 7.3: Efficiencies of the preselection and Level-1 and high level triggers, and electron contamination for the electro-weak backgrounds.

	QCD multi-jet events $\hat{\mathrm{p}}_{\mathrm{T}}$ (GeV/c)			
	120-170	170-230	230-300	300-380
Preselection	-	-	-	-
L1 trigger ($\mathrm{E_T} > 93$ GeV)	46.8%	48.5%	43.9%	26.5%
$\mathrm{E_T^{miss}} > 60$ GeV	1.5%	4.3%	8.6%	30.0%
Isolation, $\mathrm{p_T} > 25$ GeV/c	3.4%	3.6%	4.4%	3.3%
Total efficiency	2.3×10^{-4}	7.5×10^{-4}	1.7×10^{-3}	2.7×10^{-3}
Electron contamination	-	-	-	-

Table 7.4: Efficiencies of the preselection and Level-1 and high level triggers, and electron contamination for the QCD background.

and muons and which contained thus more than one neutrinos. After the isolated lepton veto, a cut was placed on the $\mathrm{E_T^{miss}}$ measurement. Then a series of cuts were applied to the events for the reliable identification the hard τ jet. The associated production mode of the charged Higgs boson was exploited by requiring tagging of one associated b jet and by reconstructing the mass of the associated top quark. A veto was applied to all other central jets which were not involved in the τ-jet identification or the top mass reconstruction. Finally, the transverse mass of the charged Higgs boson was reconstructed from the τ jet and $\mathrm{E_T^{miss}}$, and a cut based on the transverse mass window or the angular separation of the τ jet, and the $\mathrm{E_T^{miss}}$ directions was applied to the events.

The offline selection criteria are described in more detail in the following subsections.

7.3.1 Jet and track reconstruction

The hadronic jets were reconstructed in a jet cone of $\mathrm{R_{jet}} = 0.50$ with the iterative cone algorithm described in Section 4.3.1.1. The τ-jet candidates were reconstructed in a cone of $\Delta\mathrm{R} = 0.40$ around the direction of the Level-1 τ jet direction. The MC based energy corrections described in Section 4.3.1.5 were applied to correct the energy of τ jets. The energy of the hadronic jets with $\mathrm{E_T^{raw}} > 20$ GeV was calibrated with correction factors obtained from a simulation of QCD γ+jet events [132].

The tracks in the events were reconstructed with the Kalman track finder and smoother described in

Sections 4.3.2.1 and 4.3.2.2. The primary vertex was reconstructed by summing the transverse momenta of the tracks associated to each primary vertex candidate and by selecting the primary vertex candidate which produced the largest sum.

7.3.2 Isolated lepton veto

Even though the $\mathrm{H}^{\pm} \to \tau^{\pm}\nu_{\tau}$ decays were forced to decay hadronically, the W's of the associated top quark decays can produce electrons or muons in the signal samples. Furthermore, the W's occurring in the electro-weak backgrounds can produce electrons or muons. The presence of such leptons in the final state may deteriorate the $\mathrm{E}_{\mathrm{T}}^{\mathrm{miss}}$ measurement, because of their accompanying neutrino. Therefore, a veto was applied on the isolated electrons and muons.

The reconstructed electrons and muons were required to be isolated by demanding, that no reconstructed tracks with $\mathrm{p}_{\mathrm{T}} > 1.0$ GeV/c were found in a cone of $\Delta\mathrm{R} = 0.40$ around the lepton direction. If an isolated electron or muon with $\mathrm{p}_{\mathrm{T}} > 15$ GeV/c was found, the event was discarded. The hadron contamination to the electron selection was further reduced by matching tracks to the ECAL clusters [133]. The reconstructed track was identified as electron, if the matched ECAL cluster contained more than 60% of the energy carried by the track, and if the energy deposited in the HCAL cell behind the matched ECAL cluster was less than 10% of the energy of the matched ECAL cluster.

7.3.3 Missing E_{T} measurement

The missing transverse energy in the event was reconstructed with the Type-I $\mathrm{E}_{\mathrm{T}}^{\mathrm{miss}}$ corrections as described in Section 5.3. The Type-I corrections were observed to overcorrect the $\mathrm{E}_{\mathrm{T}}^{\mathrm{miss}}$ scale by 14% for signal with $\mathrm{m}_{\mathrm{H}^{\pm}} = 200$ GeV/c^2. The overcorrection was found to be $\sim$5% for signal with $\mathrm{m}_{\mathrm{H}^{\pm}} = 300$ and 400 GeV/c^2. In order to suppress efficiently QCD multi-jet events, the corrected missing transverse energy was required to exceed 100 GeV in the event.

7.3.4 Identification of the τ jet

A standard set of τ-jet identification criteria based on the methods described in Section 5.1 were applied to the events for the reliable identification of the τ jet coming from the $\mathrm{H}^{\pm} \to \tau^{\pm}\nu_{\tau}$ decay. The corrected transverse energy of the τ-jet candidate was required to exceed 100 GeV, which is somewhat higher than the required Level-1 single τ trigger threshold. The direction of the τ-jet candidate was required to be within $|\eta| < 2.5$.

The τ helicity correlations described in Sections 2.3.2 and 5.1.12 were exploited by requiring the leading track to carry at least 80% of the visible τ-jet energy. The distribution of the R_{τ} variable for τ jets matched to the $\mathrm{H}^{\pm} \to \tau^{\pm}\nu_{\tau}$ and $\mathrm{W}^{\pm} \to \tau^{\pm}\nu_{\tau}$ decays are shown in Fig. 7.1 for the different hadronic

τ decay modes. Events, in which the charged pion of the τ decay was found to carry only little of the visible τ jet energy, were found to be suppressed too much by the single τ trigger in order to be usable. The requirement $\mathrm{R}_{\tau} > 0.80$ effectively required the $\mathrm{p_T}$ of the leading track to exceed 80 GeV/c, which is well above the cut required at the single τ trigger on the leading track $\mathrm{p_T}$. The same R_{τ} cut was used both for the one- and three-prong final states, which was found to suppress most of the three-prong τ-jets.

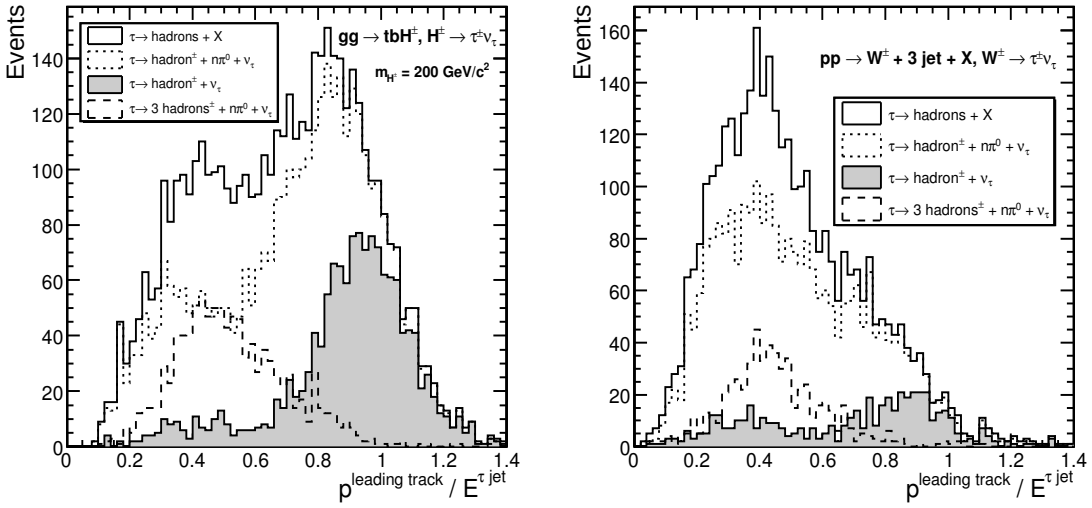

Figure 7.1: The distribution of the R_{τ} variable for τ-jet candidates matched to the $\mathrm{H}^{\pm} \to \tau^{\pm}\nu_{\tau}$ decays with $\mathrm{m_{H^{\pm}}} = 200$ GeV/c^2 (left) and matched to the $\mathrm{W}^{\pm} \to \tau^{\pm}\nu_{\tau}$ decays from the W+3 jet background (right). The dotted and dashed lines correspond to the one- and three-prong τ decays, respectively, and the filled area highlights the $\tau^{\pm} \to \pi^{\pm}\nu_{\tau}$ decay mode. The solid line is the sum of one- and three-prong τ decays.

The standard charged track isolation described in Section 5.1.6 was applied with jet matching, isolation, and signal cone sizes of $\mathrm{R_m} = 0.10$, $\mathrm{R_m} = 0.04$, and $\mathrm{R_i} = 0.40$, respectively. Tracks with $\mathrm{p_T} > 1.0$ GeV/c and with transverse impact parameter less than 0.3 mm were counted in the signal cone and in the isolation annulus. One or three tracks were required in the signal cone and no tracks were allowed in the isolation annulus. To ensure good track measurement, the leading track was required to consist of at least ten hits.

Electromagnetic isolation described in Section 5.1.5 was applied requiring, that no ECAL cell with $\mathrm{E_T} > 2.0$ GeV was found in an annulus between $\Delta\mathrm{R} = 0.10$ and $\Delta\mathrm{R} = 0.40$ around the leading track direction.

In order to further suppress the $\tau \to \mathrm{e}\nu_{\mathrm{e}}\nu_{\tau}$ decays, which survived the preceding selections, the HCAL cell with the highest transverse energy deposition in a radius of $\Delta\mathrm{R} = 0.15$ around the leading track was required to exceed 2 GeV as described in Section 5.1.10. A cut on impact parameter significance described in Section 5.1.7 was tried out, but found not to improve the background separation.

The efficiencies of the τ-jet identification criteria are shown in Tables 7.5 and 7.6 for the signal and background samples, respectively. The efficiencies of the signal and electroweak samples in the tables

are calculated relative to passing the trigger, the primary vertex reconstruction, isolated lepton veto, and the cut on the $\mathrm{E_T^{miss}}$ measurement. A preselection was performed on the QCD multi-jet sample by selecting events with at least one jet with $\mathrm{E_T} > 100$ GeV and by requiring the jet to contain a track with $\mathrm{p_T} > 80$ GeV/c before applying the τ-jet identification selections. The signal efficiencies for the τ-jet identification were found to be 14-21% with high purity. The efficiency for the electroweak and QCD multi-jet backgrounds were found to be 4-11% and 1.7%, respectively. The relatively low background rejection at the stage of τ-jet identification is explained by the fact, that the preselections have already imposed loose τ-jet identification cuts. Over 95% of the electroweak background events, which passed the τ-jet identification criteria, were found to contain a genuine τ jet coming from the $\mathrm{W}^{\pm} \to \tau^{\pm}\nu_{\tau}$ decay.

	$\mathrm{m_{H^{\pm}}}$=200 GeV/c^2	$\mathrm{m_{H^{\pm}}}$=300 GeV/c^2	$\mathrm{m_{H^{\pm}}}$=400 GeV/c^2
Selection efficiencies			
Jet $\mathrm{E_T} > 100$ GeV	52.1%	69.1%	81.8%
$\mathrm{R_{\tau}} > 0.8$	34.2%	34.0%	34.2%
1 or 3 tracks	95.9%	95.9%	97.0%
Charged track isolation	94.9%	95.3%	93.7%
Electromagnetic isolation	95.7%	95.4%	95.8%
Electron rejection	95.5%	98.0%	98.7%
$\mathrm{IP_T^{ldg.track}} < 0.3$ mm	99.0%	99.5%	99.3%
$\mathrm{N_{hits}^{ldg.track}} \geq 10$	94.6%	97.3%	96.5%
τ-jet identification efficiency and purity			
Total efficiency	13.9%	18.6%	21.1%
Signal τ purity	98.5%	99.8%	99.8%

Table 7.5: Efficiencies of the τ-jet identification methods and τ-jet identification purity after passing the trigger, PV reconstruction, isolated lepton veto and cut on $\mathrm{E_T^{miss}}$ for the signal of $\mathrm{m_{H^{\pm}}}$=200, 300, and 400 GeV/c^2.

7.3.5 Associated top and W mass reconstruction

The associated top quark mass and the mass of the W boson produced in the top quark decay reconstructed with the linear approximation of the kinematic fit described in Section 5.4. The Gaussian widths for the top and W masses were taken to be 10 GeV/c^2 and 17 GeV/c^2, respectively. A method to scale the energy of the two jets, which were associated to the W mass reconstruction, was found to give only minimal improvement and was thus not used. A small improvement to the reconstructed top mass was found to be given by correcting the energy of the b jets which produced muons through semileptonic decays. The correction was done by adding the four-momenta of the muons with $\mathrm{p_T} > 5$ GeV/c to the jet energy. The resulting top mass distribution and its fit are shown in Fig. 7.2. The reconstructed W boson and top quark masses were required to be within mass windows of $60 < \mathrm{m_{W^{\pm}}} < 100$ GeV/c^2,

	$\mathrm{t\bar{t}}$	Wt	W+3 jets	QCD
Selection efficiencies				
Jet $\mathrm{E_T} > 100$ GeV	62.6%	67.2%	69.2%	-
$\mathrm{R}_{\tau} > 0.8$	21.2%	23.7%	13.7%	40.7%
1 or 3 tracks	88.0%	93.7%	85.1%	63.0%
Charged track isolation	89.1%	91.9%	85.6%	61.4%
Electromagnetic isolation	93.7%	94.9%	94.1%	86.6%
Electron rejection	92.4%	94.2%	84.0%	99.3%
$\mathrm{IP_T}^{\mathrm{ldg.track}} < 0.3$ mm	94.8%	98.3%	84.1%	16.2%
$\mathrm{N}_{\mathrm{hits}}^{\mathrm{ldg.track}} \geq 10$	92.9%	94.4%	95.7%	74.9%
τ identification efficiency and purity				
Total efficiency	8.1%	11.4%	4.4%	1.65%
τ's from $\mathrm{W}^{\pm} \to \tau^{\pm}\nu_{\tau}$ decays	96.0%	98.7%	95.8%	-

Table 7.6: Efficiencies of the τ-jet identification methods for the electro-weak backgrounds and QCD multi-jet events with $170 < \hat{\mathrm{p}}_{\mathrm{T}} < 380$ GeV/c. The electroweak samples have been passed through the trigger, PV reconstruction, isolated lepton veto and cut on $\mathrm{E_T^{miss}}$ prior to the τ identification. The QCD multi-jet events have been required to contain at least one jet with $\mathrm{E_T} > 100$ GeV and with a leading track with $\mathrm{p_T} > 80$ GeV/c.

and $140 < \mathrm{m_t} < 210$ GeV/c^2. These mass windows corresponded roughly to a spread of $\pm 2\sigma$ of the reconstructed mass distributions.

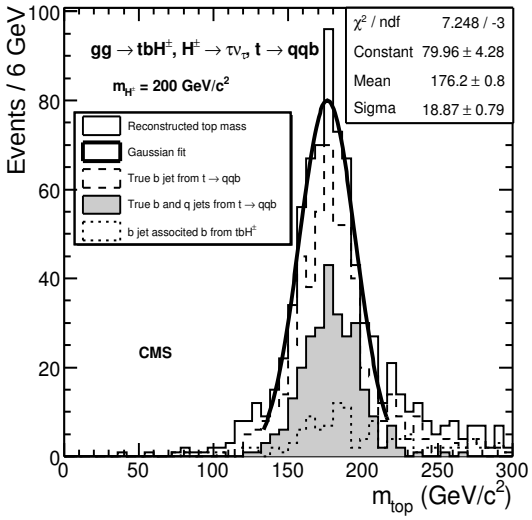

Figure 7.2: Reconstructed top mass of the associated top quark in the signal sample with $\mathrm{m_{H^{\pm}}}$ = 200 GeV/c^2 (solid lines). The dashed line and shaded area correspond to the correct b jet and correct b and q jet assignments, respectively. The dotted line shows the distribution in which the associated b jet from $\mathrm{gg} \to \mathrm{tbH^{\pm}}$ has been included in the top mass fit.

7.3.6 Identification of associated b jets

The associated charged Higgs boson production $\mathrm{gg} \to \mathrm{tbH}^{\pm}$ produces two b jets of which one comes from the associated top quark decay and the other from the production process itself. The kinematical difference between these two b jets in the event was found to be, that the b jet coming from the top decay prefers the central directions, whereas the associated b jet is preferentially emitted in the forward directions. Additionally, the b quarks coming from the top decay were found to be harder than the associated b jet in $\sim$80% of the signal events. Since the associated b jet is the hardest of the two b jets in the remaining $\sim$20% events, there was a considerable chance of misidentifying the associated b jet to be the b jet from the top decay.

The requirement of identifying both b jets was found to reduce the signal too much. Therefore, at least one b jet was required in the event. To conform with the top quark reconstruction, one of the jets associated with the top quark reconstruction was required to be identified as a b jet. The requirement, that the identified b jet would not be associated to the W reconstruction, turned out to cut too many signal events, and was hence not used.

The performance of both the track counting based and probability based b tagging algorithms, which have been described in Sections 5.2.1 and 5.2.2, respectively, was evaluated. Of these algorithms, the probabilistic b tagging algorithm was found to perform better and to yield a higher b-jet purity than the track counting algorithm. Therefore, the probabilistic b tagging algorithm was used for the identification of the b jets. The b-jet candidates were taken to be b jets, if the discriminator of the tagging algorithm exceeded 1.5, and if the transverse energy of the b-jet candidate exceeded 30 GeV.

7.3.7 Veto on additional central jets

In the signal events which have passed the selections so far, one jet has been identified as a τ jet, three jets of which one has been identified as b jet have been associated to the top mass reconstruction, and the associated b jet has possibly been identified. The associated b jet of the signal events was found to be in general much softer than the jets remaining after the top mass reconstruction in the electroweak backgrounds. Therefore, a veto was placed on events where the transverse energy of any central jet ($|\eta| < 2.5$), which did not belong to the top mass reconstruction or τ-jet identification, exceeded 25 GeV. Fake jets, which did not contain tracks, were suppressed by requiring the sum of energy carried by the tracks in the jet to be at least 10% of the jet energy.

7.3.8 Charged Higgs boson transverse mass reconstruction

Applying the conservation of the 4-momentum to the standard two-body decay of $\mathrm{H}^{\pm} \to \tau^{\pm}\nu_{\tau}$ yields

$$\mathrm{m_{H^{\pm}}}^2 = 2\mathrm{E}_{\tau}\mathrm{E}_{\nu} - \mathrm{p}_{\tau}\mathrm{p}_{\nu}\cos\phi_{\tau,\nu} + \mathrm{m_{\tau}}^2, \tag{7.1}$$

where $\Delta\phi$ is the total angle between the visible τ jet and the neutrino directions and where the mass of the neutrino has been assumed to be negligible. Since $\mathrm{p}_{\tau} \gg \mathrm{m}_{\tau}$ and since the selection cuts in the previous subsections are rather strong, equation (7.1) can be rewritten for the plane transverse to the beam axis as

$$\mathrm{m_T} = \sqrt{2\mathrm{E_T}^{\tau\,\mathrm{jet}} \times \mathrm{E_T^{miss}} \times \left(1 - \cos\Delta\phi(\tau\,\mathrm{jet}, \mathrm{E_T^{miss}})\right)}, \tag{7.2}$$

which gives an upper limit for the $\mathrm{H}^{\pm}$ candidate transverse mass when $\cos\Delta\phi$ = -1. It follows from equation (7.2) that electroweak background events in which the τ-jet and the neutrino come from the $\mathrm{W}^{\pm} \to \tau^{\pm}\nu_{\tau}$ decay acquire transverse mass values only up to $\mathrm{m_{W^{\pm}}}$. Since $\mathrm{m_{H^{\pm}}} > \mathrm{m_{W^{\pm}}}$, the $\mathrm{H}^{\pm} \to \tau^{\pm}\nu_{\tau}$ decays can be efficiently separated from $\mathrm{W}^{\pm} \to \tau^{\pm}\nu_{\tau}$ decays with on-shell W's by requiring the reconstructed transverse mass to exceed $\mathrm{m_{W^{\pm}}}$ as can be seen in Fig. 7.3. Because the number of background events with $\mathrm{W}^{\pm} \to \tau^{\pm}\nu_{\tau}$ decays is large compared to the $\mathrm{H}^{\pm} \to \tau^{\pm}\nu_{\tau}$ decays, it is possible, that some background events can acquire transverse mass beyond $\mathrm{m_{W^{\pm}}}$ as a result of the $\mathrm{E_T^{miss}}$ resolution and hence contaminate the signal region. The QCD multi-jet events can acquire transverse masses even beyond the $\mathrm{m_{H^{\pm}}}$ values, since hadronic jets may fake τ jets and/or $\mathrm{E_T^{miss}}$. Therefore, the elimination of the QCD multi-jet events with the event selections is important.

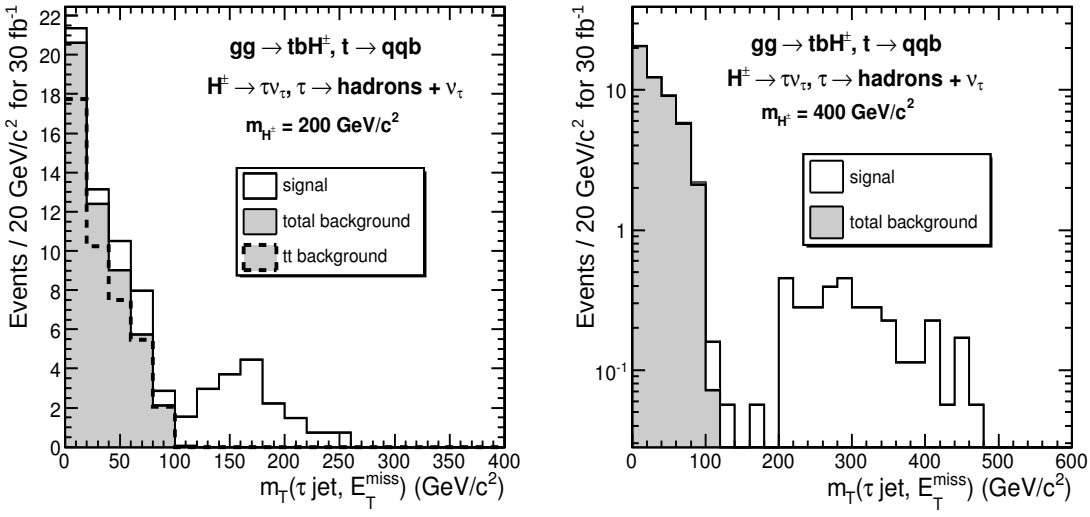

Figure 7.3: Transverse mass reconstructed from the τ jet and $\mathrm{E_T^{miss}}$ for $\mathrm{m_{H^{\pm}}}$ = 200 GeV/c^2 (left plot) and $\mathrm{m_{H^{\pm}}}$ = 400 GeV/c^2 (right plot) for the fully hadronic signal events (dark histogram) and for the total background (light histogram) for an integrated luminosity of 30 fb^{-1}. All selection cuts except the cut on the transverse mass have been applied.

Instead of the cut on the transverse mass, another possibility is to require $\Delta\phi(\tau\,\mathrm{jet}, \mathrm{E_T^{miss}})$ to exceed a certain angle instead of setting a low bound for the transverse mass. These cuts were found to be highly correlated and hence only one of them is applied.

7.4 Results

In the following, the efficiency of the trigger and offline event selections described in the Sections 7.2 and 7.3, respectively, and the corresponding number of events for an integrated luminosity of 30 fb^{-1} are summarized for the signal and the background processes. The number of events shown in the summary tables refers to the number of events expected after the corresponding selection. The efficiency of each selection was evaluated relative to the previous selection. The number of events after all selections is shown separately for the transverse mass limit and for the $\Delta\phi$ cut.

The results for the signal samples with $\mathrm{m_{H^{\pm}}}$ = 200, 300, and 400 GeV/c^2 and $\tan\beta$ = 30 are summarized in Table 7.7 for the $\mathrm{m_h^{max}}$ MSSM scenario for an integrated luminosity of 30 fb^{-1}. The number of expected signal events after all the selections was found to be (3.5±0.7)-(11.6±2.3) depending on $\mathrm{m_{H^{\pm}}}$.

	$\mathrm{m_A}$=183 GeV/c^2 $\mathrm{m_{H^{\pm}}}$=200 GeV/c^2	$\mathrm{m_A}$=289 GeV/c^2 $\mathrm{m_{H^{\pm}}}$=300 GeV/c^2	$\mathrm{m_A}$=392 GeV/c^2 $\mathrm{m_{H^{\pm}}}$=400 GeV/c^2
Cross-sections and branching ratios			
σ (fb)	1589	687	320
$\sigma \times \mathrm{BR}$ (fb)	776	118	37.7
Number of events after experimental selections			
Total events for 30 fb^{-1}	2.33×10^4	3.56×10^3	1.13×10^3
Level-1 + HLT trigger	2872 (12.3%)	864 (24.3%)	364 (32.2%)
Primary vertex found	2848 (99.2%)	856 (99.1%)	358 (98.4%)
Isolated lepton veto	2342 (82.2%)	729 (85.2%)	304 (85.0%)
$\mathrm{E_T^{miss}} > 100$ GeV	1599 (68.3%)	541 (74.2%)	245 (80.6%)
τ-jet identification	221 (13.8%)	100 (18.6%)	55.0 (22.4%)
$\geq$3 jets, $\mathrm{E_T} > 20$ GeV	171 (77.4%)	73.6 (73.4%)	39.5 (71.8%)
top mass window	109 (63.8%)	51.6 (70.1%)	27.2 (69.0%)
b tagging	46.5 (42.7%)	22.7 (44.0%)	11.0 (40.4%)
$\mathrm{E_T^{b\ jet}} > 30$ GeV	42.6 (91.7%)	19.7 (86.8%)	9.7 (88.1%)
Jet veto, $\mathrm{E_T} > 25$ GeV	15.5 (36.5%)	10.1 (51.5%)	4.0 (40.9%)
$\mathrm{E_T^{H^{\pm}}} > 50$ GeV	15.5 (100%)	8.7 (86.2%)	3.8 (94.3%)
Number of signal events with cut on $\mathrm{m_T}(\tau\ \mathrm{jet}, \mathrm{E_T^{miss}})$			
$\mathrm{m_T}(\tau\ \mathrm{jet}, \mathrm{E_T^{miss}}) > 100$ GeV/c^2	11.6±2.3 (75.1%)	8.3±1.2 (95.6%)	3.6±0.5 (94.1%)
Number of signal events with cut on $\Delta\phi(\tau\ \mathrm{jet}, \mathrm{E_T^{miss}})$			
$\Delta\phi(\tau\ \mathrm{jet}, \mathrm{E_T^{miss}}) > 60^{\mathrm{o}}$	8.3±2.0 (53.3%)	7.1±1.1 (81.4%)	3.5±0.7 (92.4%)

Table 7.7: The summary table of the event selections for the signal of $\mathrm{m_{H^{\pm}}}$=200, 300, and 400 GeV/c^2.

Since no MC event of the electroweak backgrounds passed all event selections, the selections had to be factorized to obtain estimates for the number of events after the event selections. For the $\mathrm{t\bar{t}}$ and Wt backgrounds, it was sufficient to factorize only the b tagging out of the selections. For the W+3 jet background, also the τ-jet identification without the cut on the $\mathrm{E_T}$ of the τ-jet candidate was factorized in addition to the b tagging. The results for the electroweak backgrounds are summarized in Table 7.8. The number of expected events for an integrated luminosity of 30 fb^{-1} for the electroweak backgrounds

was found to be 1.6±0.7 (0.9±0.5) after all the selections with the transverse mass limit ($\Delta\phi$ cut) applied.

	$\mathrm{t\bar{t}}$	Wt	W+3 jets
Cross-sections and branching ratios			
σ (fb)	8.40×10^{5}	6.20×10^{4}	3.55×10^{6}
$\sigma \times \mathrm{BR}$ (fb)	1.24×10^{5}	9.14×10^{3}	4.19×10^{5}
Number of events after experimental selections			
Total events for 30 fb^{-1}	3.7×10^{6}	2.74×10^{5}	1.26×10^{7}
Preselection	1.93×10^{5}	7.02×10^{3}	-
Level-1 + HLT trigger	9.61×10^{3} (4.1%)	615 (8.8%)	1.26×10^{5} (1.0%)
Primary vertex found	9.59×10^{3} (99.8%)	613 (99.7%)	1.26×10^{5} (100%)
Isolated lepton veto	8.66×10^{3} (89.4%)	551 (89.9%)	1.04×10^{5} (82.5%)
$\mathrm{E_T^{miss}} > 100$ GeV	7.37×10^{3} (85.1%)	478 (86.6%)	8.02×10^{4} (77.1%)
τ-jet identification	596 (8.1%)	54.3 (11.4%)	3.53×10^{3} (4.4%)
≥ 3 jets, $\mathrm{E_T} > 20$ GeV	518 (87.0%)	31.3 (57.6%)	2.81×10^{3} (79.8%)
top mass window	365 (70.4%)	21.2 (67.7%)	1.11×10^{3} (39.4%)
b tagging	174 (47.7%)	10.2 (48.1%)	45.7 (4.1%)
$\mathrm{E_T^{b\ jet}}$ >30 GeV	158 (90.6%)	9.1 (89.2%)	34.3 (75.1%)
Jet veto, $\mathrm{E_T}$ >25 GeV	44.3 (28.1%)	7.1 (78.0%)	5.9 (17.2%)
$\mathrm{E_T^{H^{\pm}}}$ >50 GeV	43.0 (97.1%)	7.0 (98.6%)	5.8 (98.3%)
Number of events with cut on $\mathrm{m_T}(\tau\ \mathrm{jet}, \mathrm{E_T^{miss}})$			
$\mathrm{m_T}(\tau\ \mathrm{jet}, \mathrm{E_T^{miss}})$ >100 GeV/c^2	0.86±0.33 (2.0%)	0.09±0.04 (1.3%)	0.60±0.60 (10.3%)
Number of events with cut on $\Delta\phi(\tau\ \mathrm{jet}, \mathrm{E_T^{miss}})$			
$\Delta\phi(\tau\ \mathrm{jet}, \mathrm{E_T^{miss}}) > 60^{\mathrm{o}}$	0.43±0.25 (1.0%)	0.03±0.02 (0.4%)	0.39±0.39 (6.7%)

Table 7.8: The summary table of the event selections for the electroweak backgrounds.

For the full analysis of the QCD multi-jet background, no trigger was applied. Instead, the jet with the highest $\mathrm{E_T}$ was required to exceed 100 GeV and it was required to contain at least one track with $\mathrm{p_T} > 80$ GeV/c. To mimic the high level trigger, the high level trigger threshold was applied to the uncorrected $\mathrm{E_T^{miss}}$ before applying the offline cut to the corrected $\mathrm{E_T^{miss}}$. After the $\mathrm{E_T^{miss}}$ cut, the events were required to pass the τ-jet identification and to contain at least three jets with $\mathrm{E_T} > 20$ GeV in addition to the τ-jet. Because of lack of statistics for the QCD multi-jet background, an assumption was made, that the QCD multi-jet events would behave similarly as the three jets accompanying the W boson in the W+3 jet events for the top quark reconstruction, b tagging and central jet veto. Therefore, the number of events obtained for the QCD multi-jet background after requiring more than three jets with $\mathrm{E_T} > 20$ GeV cut was multiplied with 2.1×10^{-4} or 1.4×10^{-4} to obtain the estimated number of events for the QCD multi-jet events after all the event selections.

The results for the QCD multi-jet background are summarized in Table 7.9. With the assumption described above, the expected number of events for the QCD multi-jet background for an integrated luminosity of 30 fb^{-1} was found to be 0.14±0.14 or 0.09±0.09 after all selections for either the cut on the transverse mass limit or on the $\Delta\phi$, respectively. The amount of QCD multi-jet events in the $120 < \hat{\mathrm{p}}_{\mathrm{T}} < 170$ GeV/c

bin was found to be suppressed effectively by the high jet $\mathrm{E_T}$ and $\mathrm{E_T^{miss}}$ cuts and therefore the analysis was limited to the $170 < \hat{\mathrm{p}}_{\mathrm{T}} < 380$ GeV/c range.

	QCD $170 < \hat{\mathrm{p}}_{\mathrm{T}} < 380$ GeV/c
Cross-sections and branching ratios	
σ (fb)	1.31×10^{8}
Number of events after experimental selections	
Total events for 30 fb^{-1}	3.94×10^{9}
Jet $\mathrm{E_T} > 100$ GeV, leading track $\mathrm{p_T} > 80$ GeV/c	2.19×10^{7} (0.56%)
Uncorrected $\mathrm{E_T^{miss}} > 60$ GeV	9.76×10^{5} (4.5%)
Corrected $\mathrm{E_T^{miss}} > 100$ GeV	1.49×10^{5} (15.3%)
τ identification	2460 (1.65%)
≥ 3 jets, $\mathrm{E_T} > 20$ GeV	654 (26.6%)
top mass, b tagging, jet veto, $\mathrm{E_T^{H^{\pm}}} > 50$ GeV	From W+3 jets, see text
Number of events with cut on $\mathrm{m_T}(\tau\ \mathrm{jet}, \mathrm{E_T^{miss}})$	
$\mathrm{m_T}(\tau\ \mathrm{jet}, \mathrm{E_T^{miss}}) > 100$ GeV/c^2	0.14 ± 0.14
Number of events with cut on $\Delta\phi(\tau\ \mathrm{jet}, \mathrm{E_T^{miss}})$	
$\Delta\phi(\tau\ \mathrm{jet}, \mathrm{E_T^{miss}}) > 60^{\mathrm{o}}$	0.09 ± 0.09

Table 7.9: The summary table of the selections for the QCD multi-jet background.

7.5 Systematics uncertainties and evaluation of the backgrounds from data

In the following subsections, the systematic uncertainties affecting the charged Higgs boson mass distribution are discussed and methods are presented for the evaluation of the QCD multi-jet and electroweak backgrounds from the data. The plans for measuring the jet $\to \tau$ fake rate and the efficiency of τ-jet identification have been discussed earlier in Section 5.1.13.

7.5.1 Systematic uncertainties

The systematic uncertainty resulting from the jet energy scale was estimated by varying the jet energy and the $\mathrm{E_T^{miss}}$ values with the expected jet energy scale uncertainties. The uncertainties were taken to be 3% and 2% for the $\mathrm{E_T^{miss}}$ cut efficiency and for the efficiency of selecting the three hadronic jets for the associated top quark reconstruction, respectively. The uncertainty of the τ-jet identification was estimated to be 8% for the τ jets in the $\mathrm{E_T}$ scales of $\mathrm{Z} \to \tau\tau$ decays [134]. For the b-tagging uncertainty, a conservative estimate of 5% was taken.

For the $\mathrm{t\bar{t}}$ background, the theoretical NLO cross-section uncertainty, which has been derived from the scale uncertainty, was taken to be 5% [135]. The parton distribution function uncertainty was taken to be 2.5% [136]. These values yield a total systematic uncertainty of 11% for the $\mathrm{t\bar{t}}$ background.

For the W+3 jets and QCD multi-jet backgrounds, the uncertainties resulti ng from the limited MC statistics strongly dominate the measurement uncertainties. Therefore the statistical uncertainties were used for these backgrounds instead of the systematic ones.

The total number of background events in the signal area $\mathrm{m_T}(\tau$ jet,$\mathrm{E_T^{miss}}) > 100$ GeV/c^2 was estimated to be 1.7$\pm$0.7 events including the systematic and MC uncertainties.

7.5.2 Measurement of the backgrounds from data

The occurrence of background events in the signal area, $\mathrm{m_T} > 100$ GeV/c^2, is mainly caused by the tail resulting from measurement uncertainties on $\mathrm{E_T^{miss}}$ in the electroweak background containing $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decays. Another reason is the possibility of hadronic jets to fake τ jets and $\mathrm{E_T^{miss}}$ in QCD multi-jet and W+3 jet events.

After the publication of the physics TDRs, a method has been developed to estimate the $\mathrm{t\bar{t}}$ and W+jets backgrounds from data. The method is based on choosing events with $\mathrm{W}^{\pm} \rightarrow \mu^{\pm}\nu_{\mu}$ decays from data with a hard muon selection ($\mathrm{E_T^{\mu}} > 100$ GeV) and with at least three hadronic jets. In these events, the reconstructed muon is replaced with a MC τ which has the same four-momentum as the reconstructed muon. The event is then re-reconstructed with the simulated τ decay before applying all the selections and reconstructing the charged Higgs boson transverse mass distribution. Since the muon can be reconstructed very precisely, the measurement uncertainty of the transverse mass is dominated by the $\mathrm{E_T^{miss}}$ uncertainty in this approach.

7.6 Discovery potential

The left plot of Fig. 7.4 shows the discovery potential for the heavy charged Higgs bosons in the $\mathrm{m_A}$,$\tan\beta$ plane in the $\mathrm{m_h^{max}}$ MSSM scenario with μ = 200 GeV/c^2. The statistical significance was calculated with Poissonian statistics [129]. In the figure, the SUSY corrections were included only in the branching fractions.

After the publication of the physics TDRs, the discovery potential with full SUSY corrections was studied in Ref. [137]. The corresponding discovery potential is shown in the right plot of Fig. 7.4 with different μ parameter values. The discovery contour was found to increase to about $\tan\beta$ = 35 at the low $\mathrm{m_{H^{\pm}}}$ mass region for μ = 200 GeV/c^2 and to $\tan\beta > 50$ for μ = 1000 GeV/c^2. Negative μ parameter values would enable the discovery at somewhat lower $\tan\beta$ values than what is shown in the left plot of Fig. 7.4.

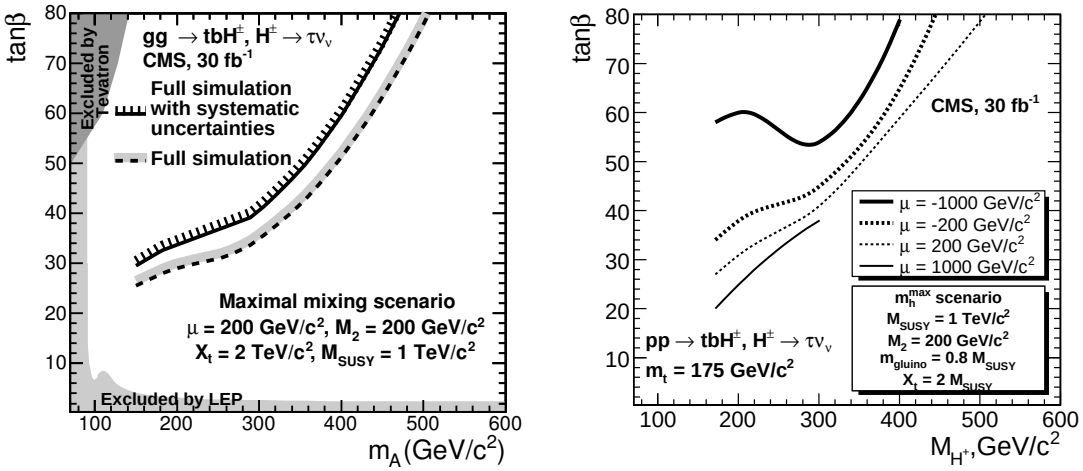

Figure 7.4: The 5σ discovery contours in the $\mathrm{m_A}$,$\tan\beta$ plane (above the curves) for $\mathrm{gg} \to \mathrm{tbH}^{\pm}$, $\mathrm{H}^{\pm} \to \tau^{\pm}\nu_{\tau}$ in the $\mathrm{m_h^{max}}$ scenario with an integrated luminosity of 30 fb^{-1} with SUSY corrections applied only to the branching fractions (left plot) from Ref. [130] and with full SUSY corrections (right plot) from Ref. [137].

7.7 Further developments after the publication of the physics TDRs

After the publication of the physics TDRs, some effects which affect the results shown in Sections 7.1-7.6 have been found. One known effect is that the simulation of the HCAL has become more realistic in CMSSW leading to some worsening in the $\mathrm{E_T^{miss}}$ resolution. Another small effect comes from the simulation of the W width including the Breit-Wigner tails in the W+jets background. The transverse mass of electroweak background events containing off-shell W's in the high end of the Breit-Wigner tail can become comparable to those of $\mathrm{m_{H^{\pm}}}$ and can hence contaminate the signal region of the transverse mass. The effect of these off-shell $\mathrm{W}^{\pm} \to \tau^{\pm}\nu_{\tau}$ decays is presently included only in the W+3/4 jet generation.

A more important effect has been found in the simulation of the $\mathrm{t\bar{t}}$ background, where the other top quark was required to produce the τ jet through the $\mathrm{W}^{\pm} \to \tau^{\pm}\nu_{\tau}$ decay and the other top quark was forced to decay fully hadronically in the full analysis presented in Sections 7.1-7.6. If the $\mathrm{t\bar{t}}$ events are not forced to decay fully hadronically, the additional neutrinos from the $\mathrm{W}^{\pm} \to \ell^{\pm}\nu_{\ell}$ decays coming from the second top quark decay can change the magnitude and the direction of the $\mathrm{E_T^{miss}}$ vector. Consequently, the reconstructed transverse mass of these events is no longer limited by the $\mathrm{m_{W^{\pm}}}$, but will instead potentially contaminate the signal region. In the full analysis, the associated $\mathrm{W}^{\pm} \to \mathrm{e}^{\pm}\nu_{\mathrm{e}}$ and $\mathrm{W}^{\pm} \to \mu^{\pm}\nu_{\mu}$ decay were suppressed with the lepton veto criterion, but the associated $\mathrm{W}^{\pm} \to \tau^{\pm}\nu_{\tau} \to \mathrm{hadr.}$ decays remained a problem.

Some new selection methods have been developed after the publication of the physics TDRs. A τ veto method has been developed to reject events with a soft second τ jet from the associated $\mathrm{W}^{\pm} \to \tau^{\pm}\nu_{\tau}$

decay to suppress the $\mathrm{E}_{\mathrm{T}}^{\mathrm{miss}}$ measurement contamination resulting from the second neutrino. Furthermore, the τ-jet identification has been reoptimized with more realistic simulation with the CMSSW software In the following, the reoptimization of the τ-jet identification and the results obtained are presented in detail.

7.7.1 Reoptimized τ-jet identification

In this section, the τ-jet identification part of the full $\mathrm{H}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ study is revisited. The selections used for the identification of the hadronically decaying τ leptons, i.e. τ jets, were re-optimized with new code applied to the simulated datasets. The aims of the τ-jet identification presented in this section are to reliably identify the signal τ jets from the $\mathrm{H}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decay, and to achieve the best possible signal separation against hadronic jets from the QCD multi-jet, $\mathrm{t\bar{t}}$ and W+jets events and against τ's from $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decays in the $\mathrm{t\bar{t}}$ and W+jets events. The τ-jet identification parameters thus obtained are to be used in a later full study of the heavy charged Higgs boson discovery potential to improve the results presented in Sections 7.1-7.6.

It was shown in Section 7.3.8, that the requirement of a high jet E_{T} and $\mathrm{E}_{\mathrm{T}}^{\mathrm{miss}}$ causes the $\mathrm{E}_{\mathrm{T}}^{\mathrm{miss}}$ to originate mainly from the $\mathrm{H}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decay. Therefore, when the charged Higgs boson candidate transverse mass is reconstructed from the τ jet and $\mathrm{E}_{\mathrm{T}}^{\mathrm{miss}}$, the $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ events obtain a transverse mass below the W mass, while the $\mathrm{H}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ events acquire transverse masses up to the simulated charged Higgs boson mass values. While this effect gives good separation of the signal τ jets from the τ's from $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decays in the $\mathrm{t\bar{t}}$ and W+jets backgrounds, events with hadronic jets faking τ jets with sufficient $\mathrm{E}_{\mathrm{T}}^{\mathrm{miss}}$ can obtain transverse masses within the signal region. Since the hadronic jets might contaminate the transverse mass spectrum and since the cross-section of background events containing hadronic jets is very large compared to the signal cross-section ($\sigma(\mathrm{hadr.\ jets}, \hat{\mathrm{p}}_{\mathrm{T}} > 80$ GeV/c / $\sigma(\mathrm{H}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau} \rightarrow \mathrm{hadr.}) \gtrsim 10^6$), it is imperative to optimize the τ-jet identification against the hadronic jets. Since the QCD multi-jet events are the most abundant source of hadronic jets in the backgrounds, the optimization of the τ-jet identification was carried out against the QCD multi-jet event samples.

However, it is not possible to completely ignore the $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decays of the electroweak backgrounds, since the very off-shell W's, which are produced in the tail of the Breit-Wigner distribution, threaten to contaminate the signal region of the charged Higgs boson transverse mass spectrum. Additionally, the associated top quark of the heavy charged Higgs production may produce a $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decay which can fake a $\mathrm{H}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decay. The only major asset against the τ jets from the $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decays is the spin-parity differences of the W and the $\mathrm{H}^{\pm}$, which cause different helicity correlations for the τ jets that arise. Therefore, the selection method based on helicity correlations is optimized against the $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decays from the $\mathrm{t\bar{t}}$ and W+jets event samples. If it is possible, a cut value that is optimal against both the τ jets from $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decays and the hadronic jets will be chosen. Additionally, electrons from the $\mathrm{W}^{\pm} \rightarrow \mathrm{e}^{\pm}\nu_{\mathrm{e}}$ and $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau} \rightarrow \mathrm{e}\nu_{\mathrm{e}}\nu_{\tau}$ decays can be properly rejected

in the one-prong final state.

The limited statistics of the QCD multi-jet event samples prohibit the full analysis of τ-jet identification together with the $\mathrm{E}_{\mathrm{T}}^{\mathrm{miss}}$ measurement, b-tagging, and associated top and W mass reconstruction. These other selections are relatively uncorrelated with the τ-jet identification, except for the cut on the τ-jet energy, and hence they are to be analyzed in a future study.

In the following, the data samples used for the re-optimization are described in Section 7.7.1.1 and the reconstruction of the events is covered in Section 7.7.1.2. Section 7.7.1.3 explains the method used for optimizing the τ identification. Finally, the selections used for the τ-jet identification in the one- and three-prong final states and the results obtained are presented in Sections 7.7.1.4 and 7.7.1.5, respectively.

7.7.1.1 Simulation of the event samples

The cross-sections and event sample sizes for the signal and background samples are summarized in Table 7.10. The events were simulated in p-p collisions with $\sqrt{s}$ = 14 TeV. No pile-up was simulated.

The signal events were generated with the process $\mathrm{gg} \rightarrow \mathrm{tbH}^{\pm}$ for three charged Higgs boson mass values of 217, 312, and 409 GeV/c^2 ($\mathrm{m_A}$ = 200, 300, and 400 GeV/c^2, respectively) and for $\tan\beta$ = 30. The PYTHIA Monte Carlo generator [69, 70] with the CTEQ5L [138] structure functions was used. The production cross-sections were normalized to the NLO results of Refs. [139, 131]. The $\mathrm{H}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ branching fractions were calculated in the $\mathrm{m_h^{max}}$ SUSY scenario [39, 127] with FeynHiggs version 2.3.2 [40, 42, 123]. The τ decays were generated with TAUOLA [74] to simulate the polarized τ decays. All τ decay modes and all decay modes of the associated top quark were generated. The τ decays in the signal sample were selected to contain only the hadronic final states based on generator level information.

The $\mathrm{t\bar{t}}$+0 jets, $\mathrm{t\bar{t}}$+1 jet were generated with ALPGEN [140] without significant jet activity setting the cut $\mathrm{E_T} > 70$ GeV for the associated jets. The cross-sections for these samples were derived from the inclusive NLO $\mathrm{t\bar{t}}$ cross-section of 840 pb. The inclusive notation '$\mathrm{t\bar{t}}$' is used later to refer to both the $\mathrm{t\bar{t}}$+0 jets and $\mathrm{t\bar{t}}$+1 jet samples.

Also the W+3 jets and W+4 jets backgrounds were generated with ALPGEN. Of the W+jets backgrounds, only the W+3/4 jets processes contribute and were studied, because top mass reconstruction is mandatory for the background suppression. The kinematics of the W+3/4 jets events is different in the region of the $\mathrm{m_{W^{\pm}}}$ and in the large mass part of the Breit-Wigner tail of $\mathrm{m_{W^{\pm}}}$, where the events become signal-like. Therefore, these backgrounds were generated in two $\mathrm{m_{W^{\pm}}}$ bins, $\mathrm{m_{W^{\pm}}} < 150$ GeV/c^2 (on-shell peak) and $150 < \mathrm{m_{W^{\pm}}} < 500$ GeV/c^2 (off-shell tail), to obtain sufficient statistics in the important tail region. Like for the signal samples, the TAUOLA package was used to generate the τ decays in the $\mathrm{t\bar{t}}$+0 jets, $\mathrm{t\bar{t}}$+1 jet, W+3 jets and W+4 jets backgrounds. The notation 'W+jets' is used to refer to both the W+3 jets and W+4 jets samples regardless of the W mass.

The QCD multi-jet event background was generated with PYTHIA in the the $\hat{\mathrm{p}}_{\mathrm{T}}$ bins of $80 < \hat{\mathrm{p}}_{\mathrm{T}} < 120$ GeV/c,

$120 < \hat{\mathrm{p}}_{\mathrm{T}} < 170$ GeV/c, and $170 < \hat{\mathrm{p}}_{\mathrm{T}} < 230$ GeV/c.

Process	σ (pb)	BR($\mathrm{H}^{\pm} \to \tau^{\pm}\nu_{\tau} \to$hadr.)	MC events
$\mathrm{gg} \to \mathrm{tbH}^{\pm}$, $\mathrm{m_{H^{\pm}}}$=217 GeV/c^2	1.589 (NLO)	0.750×0.648	87500
$\mathrm{gg} \to \mathrm{tbH}^{\pm}$, $\mathrm{m_{H^{\pm}}}$=312 GeV/c^2	0.687 (NLO)	0.264×0.648	83796
$\mathrm{gg} \to \mathrm{tbH}^{\pm}$, $\mathrm{m_{H^{\pm}}}$=409 GeV/c^2	0.320 (NLO)	0.181×0.648	78783
QCD, $80 < \hat{\mathrm{p}}_{\mathrm{T}} < 120$ GeV/c	3.08×10^{6} (LO)	-	1259257
QCD, $120 < \hat{\mathrm{p}}_{\mathrm{T}} < 170$ GeV/c	4.94×10^{5} (LO)	-	1243421
QCD, $170 < \hat{\mathrm{p}}_{\mathrm{T}} < 230$ GeV/c	1.01×10^{5} (LO)	-	929870
$\mathrm{t\bar{t}}$ + 0 jets	621.5 (NLO)	-	1406142
$\mathrm{t\bar{t}}$ + 1 jet	177.0 (NLO)	-	397148
W+3 jets, $0 < \mathrm{m_{W^{\pm}}} < 150$ GeV/c^2	587.0 (LO)	-	716202
W+4 jets, $0 < \mathrm{m_{W^{\pm}}} < 150$ GeV/c^2	123.8 (LO)	-	486802
W+3 jets, $150 < \mathrm{m_{W^{\pm}}} < 500$ GeV/c^2,	1.06 (LO)	-	32290
W+4 jets, $150 < \mathrm{m_{W^{\pm}}} < 500$ GeV/c^2	0.20 (LO)	-	44482

Table 7.10: Cross-sections and size of the MC event sample for $\mathrm{gg} \to \mathrm{tbH}^{\pm}$ signal events at tanβ = 30, $\mathrm{t\bar{t}}$+0 jets, $\mathrm{t\bar{t}}$+ 1 jet, W+3 jets, W+4 jets and QCD multi-jet backgrounds. The branching fractions for $\mathrm{H}^{\pm} \to \tau^{\pm}\nu_{\tau}$ are shown for the signal samples as a function of $\mathrm{m_{H^{\pm}}}$.

7.7.1.2 Reconstruction of the events

The CMS software version 1.6.12 [68] was used to reconstruct the events. The trigger for the heavy charged Higgs bosons has been studied in detail in Refs. [61, 62]. In this study, all single τ triggers included a cut based on $\mathrm{E_T^{miss}}$. It was, however, decided to leave the selections based on $\mathrm{E_T^{miss}}$ outside the study of τ identification, because the QCD multi-jet events are strongly suppressed already at the trigger level, and because the number of MC events would have been too small to study both $\mathrm{E_T^{miss}}$ and τ identification selections at the same time. Therefore, no trigger was applied to the samples. However, care was taken to impose selections which are not biased and which are stronger than the single τ Level-1 and high level triggers without the $\mathrm{E_T^{miss}}$ selections. The left plot of Fig. 7.5 shows, that if the jet $\mathrm{E_T}$ threshold is required to be above 100 GeV, the plateau of the emulated Level-1 single τ trigger threshold has been reached and no bias is caused to the selection. Hence, the overall τ identification selection efficiencies discussed in this study contain the single τ trigger without the $\mathrm{E_T^{miss}}$ selections.

The jets were reconstructed with the iterative cone algorithm described in Section 4.3.1.1 with a jet radius of 0.50. The reconstruction of the τ jet leads to an underestimation the τ-jet energy by about 10 GeV for $\mathrm{m_{H^{\pm}}}$ = 312 GeV/c^2. To correct this effect, energy corrections based on MC studies of τ jets described in Section 4.3.1.5 were used. The right plot of Fig. 7.5 shows a clear improvement in the jet energy scale of the τ-jet candidates for corrected jets compared to the uncorrected ones after requiring a jet energy of $\mathrm{E_T} > 80$ GeV. The figure contains both one- and three-prong final states. The differences in jet resolution between these decay modes were found to be minimal.

The charged tracks inside the jet cones were reconstructed with the iterative tracking method, which

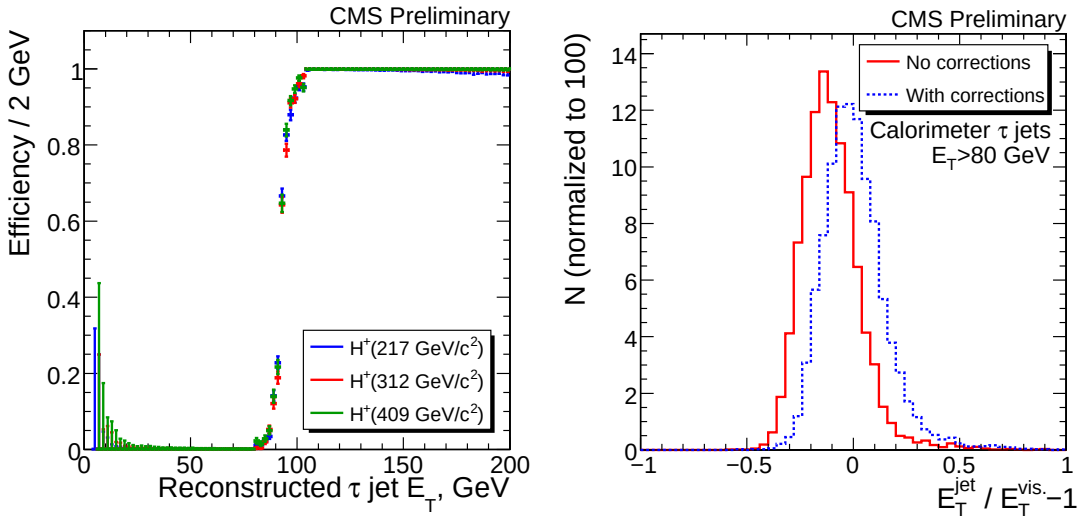

Figure 7.5: Left: Efficiency of energy corrected calorimeter τ jets passing an $E_T > 80$ GeV cut on the uncorrected calorimeter τ jet. Both jets have been matched to the MC τ with $H^{\pm}$ as ancestor. Right: Distribution of energy resolution $E_T^{jet}/E_T^{vis.}-1$ for MC matched calorimeter τ-jet candidates without (solid line) and with (dashed line) MC based τ-jet energy corrections for signal events with $m_{H^{\pm}} = 312$ GeV/c^2.

enables the reconstruction of tracks with $p_T > 0.3$ GeV/c as described in Section 4.3.2.3. At the single τ trigger level, tracks are required to have $p_T > 1.0$ GeV/c in order to be taken into account.

7.7.1.3 Optimization method

The τ-jet identification selection methods of the one- and three-prong τ decay final states were optimized and treated separately. The optimization was done by applying all the selections to the data samples except for the selection which was being optimized. The cut point of the selection being optimized was varied and for each point a Poissonian significance [141]

$$\frac{\sigma}{\sqrt{2}} = \sqrt{(N_{signal} + N_{bkg}) \ln\left(1 + \frac{N_{signal}}{N_{bkg}}\right) - N_{signal}} \tag{7.3}$$

was calculated, where N_{signal} and N_{bkg} are the number of signal and background events, respectively, for a given luminosity. Equation (7.3) was used to produce a significance curve for each of the studied background samples. For reasons of clarity, the individual background samples of same type were added together prior to the significance calculation with their corresponding normalization based on their efficiency and cross-section. It should be noted, that the summing of samples with different cross-section can lead to step-like incontinuous efficiency and significance curves in the limit of low statistics.

The optimal point of rejection was determined from the maximum point of the significance curve. As mentioned in Section 7.7.1, the QCD multi-jet background was considered as the most dangerous background and hence the maximum point of the significance curve obtained from the QCD multi-jet sample was taken as the optimal point for each cut. The R_{τ} cut was optimized also against the $t\bar{t}$ background.

After each variable was optimized, new significance curves were drawn and the variables were adjusted again correspondingly. The iteration was carried out several times over all selection variables except the cone size selections. The maxima of the significance curves were observed to become stable after a few iterations over all variables. Since the optimization was carried out by the shape of the Poissonian significance curve and not the absolute value, the optimization is independent of the integrated luminosity of the data.

The results of the optimization are portrayed in such a way (e.g. left plot of Fig. 7.6), that both the efficiency of the selection and its corresponding significance are shown as a function of the variable of the selection. The efficiency is normalized to unity for the variable and shown as a solid curve for each dataset. Furthermore, the significance is shown as a dashed curve for each background and is normalized in such a way, that the maximum significance becomes unity. The color code corresponding to each background category was kept the same in each of the optimization plots. A vertical line was used to mark the selected cut value with the shaded side of this line pointing to the direction which was being cut away. The optimization was done with the $\mathrm{m}_{\mathrm{H}^{\pm}}$ = 312 GeV/c^2 signal sample. The efficiency curves were found to be to a good approximation independent of the choice of $\mathrm{m}_{\mathrm{H}^{\pm}}$ in all cuts other than the jet E_{T} threshold (right plot of Fig. 7.6).

7.7.1.4 Identification of one-prong τ jets from $\mathrm{H}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$

The identification of one-prong final states of hadronic τ decays relies on a jet E_{T} threshold, narrow isolated decay signature in terms of charged tracks, isolated electromagnetic energy deposition, rejecting electrons, and utilizing τ helicity correlations. Additionally, standard track quality cuts were applied. These selection methods are described in the following subsections. The results are summarized in Section 7.7.1.4.6.

7.7.1.4.1 Kinematical selections

The τ-jet candidates were taken to be the jets, which exceed an optimized jet E_{T} threshold, and which had to be larger than the single tau trigger threshold of 80 GeV. The jet E_{T} threshold was observed to give a large rejection factor against the backgrounds (left plot of Fig. 7.6). When all the other selections were taken into account, the optimal jet E_{T} threshold against hadronic jets from QCD background was observed to be 119 GeV for $\mathrm{m}_{\mathrm{H}^{\pm}}$ = 312 GeV/c^2. It is evident from the left plot of Fig. 7.6, that a stronger cut than $\mathrm{E}_{\mathrm{T}} > 119$ GeV may be chosen, if necessary, to reduce more background with moderate cost in terms of signal efficiency. The signal efficiency of the jet E_{T} threshold was found to depend on the value of $\mathrm{m}_{\mathrm{H}^{\pm}}$ (right plot of Fig. 7.6), but it was observed, that a jet E_{T} threshold up to at least 150 GeV enables still the search for a charged Higgs boson with $\mathrm{m}_{\mathrm{H}^{\pm}} \sim \mathrm{m}_{\mathrm{t}}$. Since the charged Higgs boson production cross-section increases as $\mathrm{m}_{\mathrm{H}^{\pm}}$ is decreased, the net event rate remains higher for the $\mathrm{m}_{\mathrm{H}^{\pm}} \sim \mathrm{m}_{\mathrm{t}}$ than for $\mathrm{m}_{\mathrm{H}^{\pm}}$ = 312 GeV/c^2 at the same $\tan\beta$ value. It should be noted, that the jet E_{T} threshold effectively

eliminates the $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau} \rightarrow \mu\nu_{\mu}\nu_{\tau}$ decay modes from the $\mathrm{t\bar{t}}$ and W+jet backgrounds.

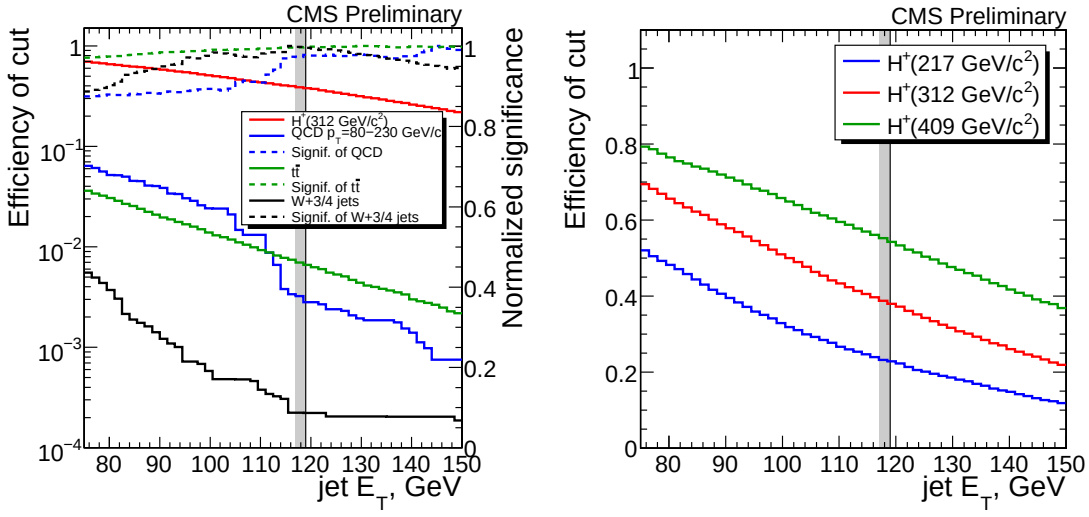

Figure 7.6: Left: Selection efficiency (solid curves) and normalized significance (dashed curves) for one-prong τ-jet candidates as a function of jet $\mathrm{E_T}$ after all other selections for finding a lower bound. Right: Selection efficiency for one-prong τ-jet candidates as a function of jet $\mathrm{E_T}$ after all other selections for the τ's matched to the $\mathrm{H}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decay with $\mathrm{m_{H^{\pm}}}$ = 217, 312, and 409 GeV/c^2.

The $\mathrm{p_T}$ of the track with the highest transverse momentum (i.e. the leading track) was required to exceed 20 GeV/c to emulate the single τ trigger. The actual $\mathrm{p_T}$ threshold applied was, however, larger, since in the charged Higgs boson searches, the cut on the leading track $\mathrm{p_T}$ can be included in the method of exploiting the helicity correlations discussed in Section 7.7.1.4.5.

The jet $|\eta|$ distribution efficiency and the corresponding significances after all the other selections is shown in Fig. 7.7. for determining an upper bound of the selection. The hadronic jets from the QCD multi-jet events are distributed to more forward directions than the τ jets from the signal and $\mathrm{t\bar{t}}$ events. The optimal cut value was found to be $|\eta| < 1.7$, which is well within the tracker acceptance region.

7.7.1.4.2 Charged track isolation

The charged track isolation algorithm described in Section 5.1.6 was applied. The tracks, which were examined in the algorithm, were required to have $\Delta\mathrm{IP_z} < 1$ mm and $\mathrm{IP_T} < 300$ μm. Only tracks, whose transverse momentum exceeded $\mathrm{p_T^{min}}$, were counted. The matching cone size of $\mathrm{R_m} = 0.10$ was chosen. The size of the isolation cone was determined by optimizing the selection significance based on the $\mathrm{p_T^{min}}$ of the tracks against the hadronic jets for different isolation cone sizes (left plot of Fig. 7.8). It was observed, that outer bound values above 0.45 give very similar performance in terms of significance for hadronic jets from QCD multi-jet events. Hence the radius $\mathrm{R_i} = 0.50$ was selected to conform with the annulus size in the trigger and to reject efficiently hadronic jets. It should, however, be noted that it may be necessary to choose a smaller isolation annulus for luminosities for which pile-up cannot be ignored. The choice of the signal cone size does not affect the one-prong selection, but it is important for the three-prong selection.

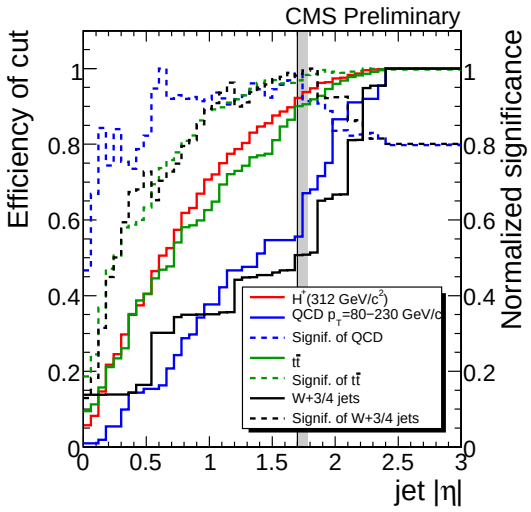

Figure 7.7: Selection efficiency (solid curves) and normalized significance (dashed curves) for one-prong τ-jet candidates as a function of jet η after all other selections for finding an upper bound.

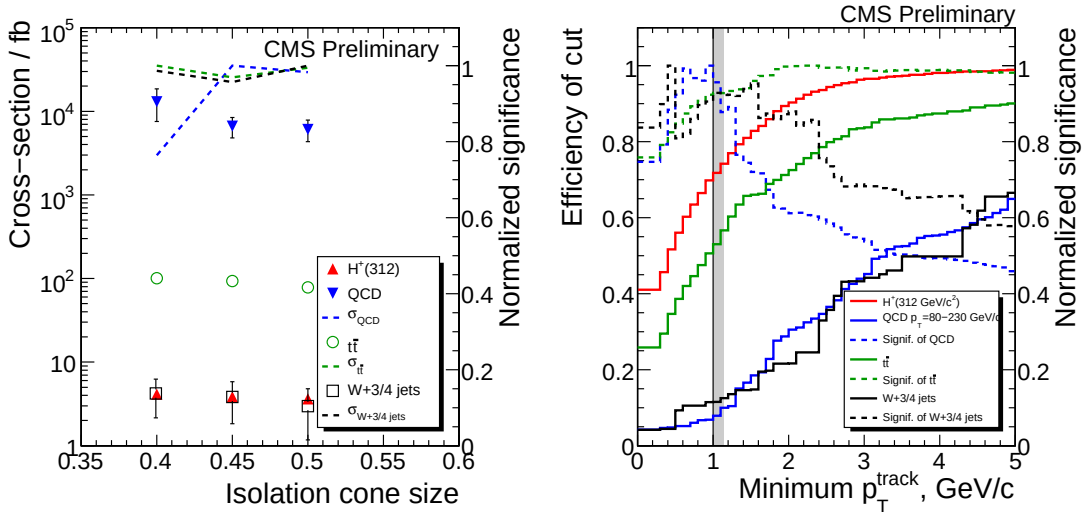

Figure 7.8: Left: Selection efficiency (points) and normalized significance (dashed curves) for one-prong τ-jet candidates as a function of charged track isolation annulus higher edge after all other selections. Right: Selection efficiency (solid curves) and normalized significance (dashed curves) for one-prong τ-jet candidates as a function of minimum track $\mathrm{p_T}$ in isolation cone after all other selections for finding a lower bound.

The minimum transverse momentum of the tracks, $\mathrm{p_T^{min}}$, which controls how soft tracks are counted in the algorithm, was varied to obtain the optimal background rejection. Lowering the $\mathrm{p_T^{min}}$ threshold causes on average more tracks to be found in the isolation cone for hadronic jets and thus lowering the threshold tightens the selection. The selection efficiency and the significance of the charged track isolation is shown in the right plot of Fig. 7.8 as a function of $\mathrm{p_T^{min}}$. The optimal performance was observed to be given by $\mathrm{p_T^{min}} = 1.0$ GeV/c, which was already set at trigger level. It was observed, that the $\mathrm{p_T^{min}}$ threshold could be lowered down to 0.5 GeV/c to suppress further the hadronic jets in the backgrounds, but the cost would be a decrease of 20-30% units in signal efficiency. Additionally, the charged track was required

to have at least eight hits and the track fit was required to fulfill normalized $\chi^2_{\mathrm{ldg.track}} < 10$. These track quality cuts yielded only minimal background rejection, but they ensured a good quality of the charged track.

7.7.1.4.3 Electromagnetic isolation

The algorithm for electromagnetic isolation described in Section 5.1.5 was applied. The signal cone and the isolation annulus were applied around the leading track momentum direction at the production vertex. The outer bound of the isolation annulus was taken to be 0.50, which is the same value as the one used for the charged track isolation. The size of the signal cone was determined by optimizing the selection significance based on the upper limit of $\mathrm{P}_{\mathrm{isol}}$ against hadronic jets from QCD multi-jet events for different signal cone sizes. It can be seen in the left plot of Fig. 7.9, that the optimal rejection was achieved, when the signal cone size of 0.10 was selected. The corresponding efficiency and significance curves for the signal cone size of 0.10 are shown in the right plot of Fig. 7.9 as a function of $\mathrm{P}_{\mathrm{isol}}$. The optimal upper limit for the electromagnetic energy in the isolation annulus, $\mathrm{P}_{\mathrm{isol}}$, was determined to be 1.8 GeV.

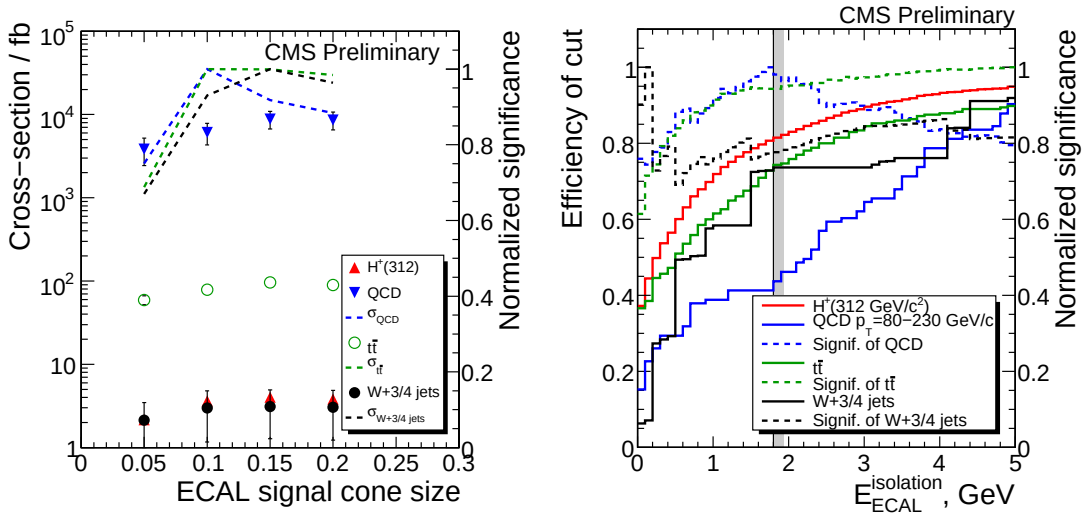

Figure 7.9: Left: Selection efficiency (points) and normalized significance (dashed curves) for one-prong τ-jet candidates as a function of electromagnetic isolation annulus inner bound after all other selections. Right: Selection efficiency (solid curves) and normalized significance (dashed curves) for one-prong τ-jet candidates as a function $\mathrm{P}_{\mathrm{isol}}$ after all other selections for finding a lower bound.

7.7.1.4.4 Electron rejection

The τ's coming from the $\mathrm{H}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decay were required to decay hadronically and do thus not contain any electrons. The amount of electrons in hadronic jets is small. The main sources of isolated electrons are thus the $\mathrm{W}^{\pm} \rightarrow \mathrm{e}^{\pm}\nu_{\mathrm{e}}$ and $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau} \rightarrow \mathrm{e}\nu_{\mathrm{e}}\nu_{\tau}$ decays in the signal, $\mathrm{t\bar{t}}$, and W+jets events. To further suppress electrons from these sources, the hadron calorimeter energy deposition was matched to the momentum of the leading track as described in Section 5.1.10. The matching was done with the

variable $\mathrm{E}_{\mathrm{T}}^{\mathrm{HCAL}}/\mathrm{p}_{\mathrm{T}}^{\mathrm{track}} - 1 > \mathrm{R}_{\mathrm{e}}$. To conform with the charged track and electromagnetic isolation annulus sizes, the hadronic energy deposition was calculated in a cone of 0.50 around the leading track momentum direction. For leading tracks, which are reconstructed electrons, the ratio is expected to be very small and these electrons may be removed by setting a lower bound on this variable. The resulting efficiency and significance curves are shown in Fig. 7.10 from which the optimal lower bound of R_{e} = -0.90 can be determined. This electron rejection is observed to suppress the electron contamination by a factor of $\sim$20 for the $\mathrm{t\bar{t}}$ and W+jets backgrounds.

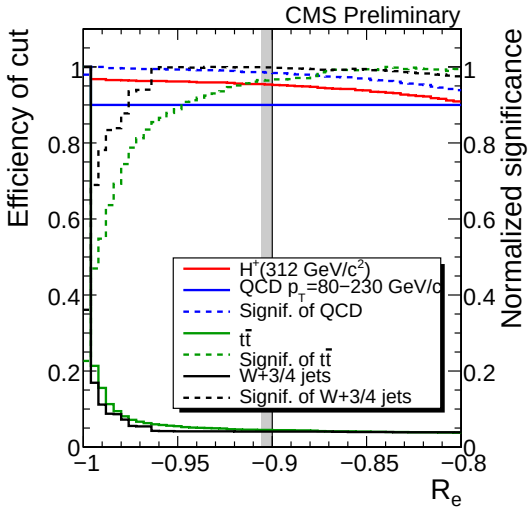

Figure 7.10: Selection efficiency (solid curves) and normalized significance (dashed curves) for one-prong τ jet candidates as a function of $\mathrm{E}_{\mathrm{HCAL}}/\mathrm{p}_{\mathrm{T}}^{\mathrm{ldg.track}}$ after all other selections for finding a lower bound R_{e}.

7.7.1.4.5 Helicity correlations

The τ helicity correlations described in Sections 2.3.2.1 and 5.1.12 were used to separate signal τ jets from τ jets coming from the $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decays. The efficiency of the helicity correlation selection after all other cuts and the corresponding significance curves are shown in the left plot of Fig. 7.11 as a function of the R_{τ} variable for τ jets from $\mathrm{H}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ and $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decays. It can be seen in the left plot of Fig. 7.11, that the optimal separation of MC matched signal τ jets against the MC matched τ jets from $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decays was achieved, when a cut $\mathrm{R}_{\tau} > 0.80$ was required. The right plot of Fig. 7.11 shows the efficiency and the corresponding significance curves of the helicity correlations for the full event samples including hadronic jets in addition to τ jets. The optimal cut point against all types of backgrounds was found to be $\mathrm{R}_{\tau} > 0.80$.

The requirement of the leading track carrying majority of the visible τ jet energy does cut a lot of the signal, as only about half of the $\rho_{\mathrm{L}}^{\pm}$ and a small part of $\mathrm{a}_{1\mathrm{L}}^{\pm}$ are selected. In an effort to recover some of the lost signal events, the selections were re-optimized separately for the region $\mathrm{R}_{\tau} < 0.4$. A part of the signal events in this region are irrecoverably cut away by the threshold of 20 GeV/c on the p_{T} of the

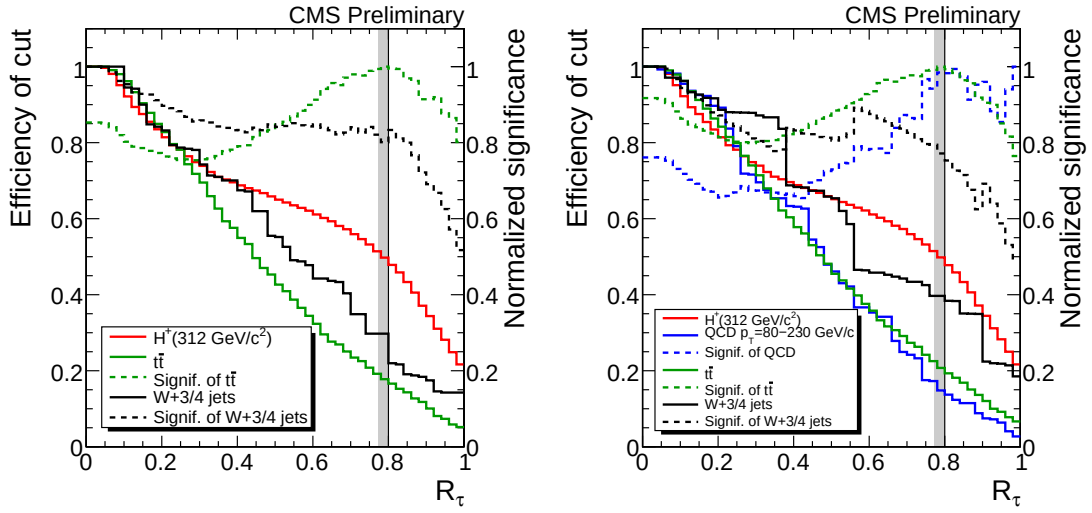

Figure 7.11: Selection efficiency (solid curves) and normalized significance (dashed curves) for MC matched hadronic one-prong τ decays (left) and one-prong τ-jet candidates (right) as a function of R_τ after all other selections for finding a lower bound.

leading track applied at trigger level. It was observed, that only minor rejection power could be achieved against the backgrounds with the R_τ variable. Consequently the signal was found to be overwhelmed in the region $R_\tau < 0.4$ by the QCD multi-jet background, even though the jet E_T threshold was increased to as high as 138 GeV.

Since it was not possible to reconstruct the neutral pions in this study, the method of energy sharing between the pions was not studied for the one-prong τ decays. Future studies with the particle flow method will enable the exploiting of such method.

7.7.1.4.6 Results for one-prong τ identification

The results for the one-prong τ identification are summarized in Tables 7.11-7.14. In the tables, each row displays the cross-section in fb for a data sample after the mentioned selections, which have been described in the previous subsections. The absolute reduction in the event rate by a certain selection method may differ from the efficiencies in the optimization plots, because of the correlations between the selection methods. Since at least as tight cuts have been demanded as in the default τ trigger, the overall τ selection efficiency includes the τ trigger without the E_T^{miss} part for the one-prong final state. The errors shown are statistical errors.

Table 7.11 shows the cross-sections after each selection criterion and the total τ selection efficiency for the signal events with $m_{H^{\pm}}$ = 217 GeV/c^2, 312 GeV/c^2, and 409 GeV/c^2. For the lowest charged Higgs boson mass, the jet E_T threshold causes substantial reduction in the event rate, but the highest charged Higgs boson mass is only little affected. The high E_T threshold is, however, needed to suppress efficiently the lowest $\hat{p}_T$ bins of the hadronic jets from QCD multi-jet background. An overall selection efficiency of ∼2-5% is reached for the signal samples with efficiency slightly rising with the charged Higgs boson

mass. The signal events, which have passed the τ identification, are observed to contain 0.21±0.15%, 0.30±0.13%, and 0.04±0.04% of τ's from associated $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decays for $\mathrm{m_{H^{\pm}}}$ = 217, 312, and 409 GeV/c^2, respectively. The amount of associated hadronic jets passing the τ identification in the signal events is slightly larger than the amount of τ's from the associated $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decay. Thus, the purity of the signal τ-jet identification is of the order of 98-99% for the signal samples.

Table 7.12 shows the cross-sections after each selection criterion and the total τ selection efficiency for the QCD multi-jet events. The cut on the jet $\mathrm{E_T}$ reduces efficiently the QCD multi-jet events in the lowest $\hat{\mathrm{p}}_{\mathrm{T}}$ bins. The charged track isolation is also effective with increasing rejection as a function of $\hat{\mathrm{p}}_{\mathrm{T}}$, since track multiplicity grows as a function of $\hat{\mathrm{p}}_{\mathrm{T}}$. The electromagnetic isolation and the cut on the R_{τ} variable give the next best rejection. Also here, the rejection increases as a function of $\hat{\mathrm{p}}_{\mathrm{T}}$, since the amount of π^0's produced in the hadronization of the jet and the energy carried by them increases as a function of $\hat{\mathrm{p}}_{\mathrm{T}}$. The overall rejection is found to be at least 10^6 for the $\hat{\mathrm{p}}_{\mathrm{T}}$ = 80-120 GeV/c bin and $\sim 10^5$ for $\hat{\mathrm{p}}_{\mathrm{T}}$ = 120-230 GeV/c.

The cross-sections after each selection and the overall τ selection efficiency for $\mathrm{t\bar{t}}$ events are shown in the first two columns of Table 7.13. Since these events contain b-jets and since the W can produce hadronic jets, the selections, which are important against the QCD multi-jet background, are also working against the $\mathrm{t\bar{t}}$ background. The electron rejection cuts efficiently $\mathrm{t\bar{t}}$ events, which contain electrons through the $\mathrm{W}^{\pm} \rightarrow \mathrm{e}^{\pm}\nu_{\mathrm{e}}$ and $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau} \rightarrow \mathrm{e}\nu_{\mathrm{e}}\nu_{\tau}$ decays. The $\mathrm{E_T}$ threshold benefits from the helicity correlations, which leads to softer τ jets in the $\mathrm{t\bar{t}}$ and W+jets events than in the signal events. Because of this effect, the $\mathrm{E_T}$ threshold is the single most efficient selection against the τ's from $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ giving a rejection factor of $\sim$47 and $\sim$24 for the $\mathrm{t\bar{t}}$+0 jets and $\mathrm{t\bar{t}}$+1 jet samples, respectively. The second most effective cut against the τ's from $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decays is the R_{τ} cut, which yields a rejection factor of $\sim$5. The overall rejection factor against the $\mathrm{t\bar{t}}$ events is observed to be $\sim 10^4$. About 65-70% of the $\mathrm{t\bar{t}}$ events surviving the τ identification were found to contain τ's from $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decays making this fraction of events irreducible with the τ selection.

The last column of Table 7.13 and the Table 7.14 show the cross-sections after each selection and the τ selection efficiency for the W+jets backgrounds. As for the $\mathrm{t\bar{t}}$ background, the signature of the W+jets contains hadronic jets, electrons and τ jets from $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decay. The most important cuts against the hadronic jets are the same as against the QCD multi-jet background. The electrons were rejected mostly with the electron rejection. Because of the helicity correlations, the $\mathrm{E_T}$ threshold rejects the τ's from $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ with a factor $\sim 2\times 10^3$ for the on-shell (peak) samples killing thus effectively the $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ final state for these samples. The overall rejection factor of τ identification is observed to be $\sim 2\times 10^5$ and $\sim 4\times 10^5$ for the on-shell W+3 jets and W+4 jets samples, respectively.

In the off-shell (tail) W+jets backgrounds, the τ's are much harder than in the on-shell samples. Thus, the $\mathrm{E_T}$ threshold is no longer as effective for τ's coming from the off-shell W decay as for the on-shell decays yielding only a rejection factor of $\sim$20. The R_{τ} cut is the second most effective selection against these backgrounds with a rejection factor of $\sim$4. The overall rejection factor against the off-shell samples is found to be $\sim 3\times 10^3$. Since the majority of the surviving events were found to contain τ

jets from $\mathrm{W}^{\pm} \to \tau^{\pm}\nu_{\tau}$ decay with similar momentum as that of the signal τ's, they are considered as irreducible background with the τ selection. Fortunately their cross-section after the τ selection is quite small compared to that of the signal samples.

7.7.1.5 Identification of three-prong τ jets from $\mathrm{H}^{\pm} \to \tau^{\pm}\nu_{\tau}$

The three-prong hadronic τ decays take place via the a_1 mode, $\tau \to \mathrm{a}_1\nu_{\tau}$. This decay mode accounts for about 15% of τ decays out of which 2/3 ($\sim$10%) comes from $\tau^{\pm} \to \pi^{\pm}\pi^{\pm}\pi^{\mp}\nu_{\tau}$ decays without the presence of neutral pions. Therefore the adopted strategy to suppress hadronic jets, which fake three-prong τ jets, was to select the decay modes without π^0's. This selection was enforced by matching the energy carried by the charged tracks to the visible jet energy.

The kinematical selections, charged track and electromagnetic isolation, and the selection based on τ helicity correlation, which were used for the one-prong τ-jet identification in Sections 7.7.1.4.1-7.7.1.4.3 and 7.7.1.4.5, were re-optimized and applied for the three-prong final state. Since there are no isolated electrons in the three-prong final state of τ decays, the electron rejection described for the one-prong final state in Section 7.7.1.4.4 was not applied in the three-prong final state. In addition, selections based on τ flight path reconstruction and invariant τ mass reconstruction were applied.

The selections for the three-prong τ-jet identification are described in the following subsections. The results are summarized in Section 7.7.1.5.8

7.7.1.5.1 Kinematical selections

The selection efficiency based on the jet $\mathrm{E_T}$ threshold and the corresponding significance curves are shown in Fig. 7.12 for three-prong τ jet candidates. It was noticed, that the amount of three-prong jets is larger than the amount of one-prong jets in the background samples and that the jet $\mathrm{E_T}$ threshold is more correlated with the other selections. Otherwise, the behavior of the selection efficiencies was observed to follow that of the one-prong τ jet candidates. The optimized jet $\mathrm{E_T}$ threshold was found to be $\mathrm{E_T} > 102$ GeV.

The optimized cut for jet η was observed to be $|\eta| < 1.8$, which is close to the optimal value found for the one-prong final state.

7.7.1.5.2 Charged track isolation

For the three-prong final state, the same charged track isolation algorithm parameters were used as those described for the one-prong final state in Section 7.7.1.4.2. Because of the boost effect, the two other charged tracks of the three-prong τ decays are accommodated inside a narrow signal cone around the leading track direction. Since the general strategy of the three-prong τ-jet identification was to select the

	$H^{\pm} \to \tau^{\pm}\nu_{\tau}$ $m_{H^{\pm}}$=217 GeV/c^2 m_A=200 GeV/c^2	$H^{\pm} \to \tau^{\pm}\nu_{\tau}$ $m_{H^{\pm}}$=312 GeV/c^2 m_A=300 GeV/c^2	$H^{\pm} \to \tau^{\pm}\nu_{\tau}$ $m_{H^{\pm}}$=409 GeV/c^2 m_A=400 GeV/c^2
Total cross-section and MC event resolution			
Cross-section (fb) / MC event	0.014	2.16×10^{-3}	7.35×10^{-4}
Cross-section (fb)	1.19×10^{3}	181	57.9
Hadronic τ jet (fb)	772	117	37.5
Cross-section after selections and the efficiency for the selection step			
Jet $E_T > 119$ GeV	303±2 fb (40.1±0.3 %)	61.0±0.4 fb (53.0±0.3 %)	23.7±0.1 fb (64.0±0.4 %)
Jet $\|\eta\| < 1.7$	273±2 fb (90.2±0.6 %)	55.7±0.3 fb (91.3±0.6 %)	21.9±0.1 fb (92.5±0.5 %)
Leading track $p_T > 20$ GeV/c	234±2 fb (85.9±0.7 %)	48.8±0.3 fb (87.6±0.6 %)	19.5±0.1 fb (88.9±0.6 %)
Ch. track isolation,$p_T^{track} > 1.0$ GeV/c	66.9±1.0 fb (28.6±0.4 %)	17.9±0.2 fb (36.8±0.4 %)	8.23±0.08 fb (42.3±0.4 %)
1 track in signal cone	48.9±0.8 fb (73.1±1.2 %)	12.7±0.2 fb (70.6±0.9 %)	5.84±0.07 fb (70.9±0.8 %)
Leading track quality cuts	47.2±0.8 fb (96.4±1.7 %)	12.3±0.2 fb (97.5±1.3 %)	5.68±0.07 fb (97.3±1.1 %)
Em. isolation $E_T^{ECAL} < 1.8$ GeV	32.5±0.7 fb (68.9±1.4 %)	8.57±0.14 fb (69.5±1.1 %)	3.92±0.05 fb (69.1±1.0 %)
Electron rejection $R_e > -0.90$	28.0±0.6 fb (86.3±1.9 %)	7.67±0.13 fb (89.5±1.5 %)	3.55±0.05 fb (90.6±1.3 %)
$R_\tau > 0.8$	13.4±0.4 fb (47.6±1.5 %)	3.71±0.09 fb (48.4±1.2 %)	1.73±0.04 fb (48.8±1.0 %)
τ-jet identification efficiency and purity			
Efficiency/event	1.73±0.06 %	3.17±0.08 %	4.62±0.10 %
Signal τ purity	99.5±0.2 %	98.8±0.3 %	99.6±0.1 %
τ's from associated $W^{\pm} \to \tau^{\pm}\nu_{\tau}$	0.21±0.15 %	0.30±0.13 %	0.04±0.04 %

Table 7.11: Cross-sections after selections and the τ identification efficiency for the signal events with $m_{H^{\pm}}$ = 217-409 GeV/c^2 in the one-prong final state. The errors represent statistical uncertainty.

	QCD $\hat{p}_T$= 80-120 GeV/c	QCD $\hat{p}_T$= 120-170 GeV/c	QCD $\hat{p}_T$= 170-230 GeV/c
Total cross-section and MC event resolution			
Cross-section (fb) / MC event	2.45×10^{3}	397	109
Cross-section (fb)	3.08×10^{9}	4.94×10^{8}	1.01×10^{8}
Cross-section after selections and the efficiency for the selection step			
Jet $E_T > 119$ GeV	$5.54\pm0.04\times10^{7}$ fb (2.04±0.01 %)	$1.37\pm0.01\times10^{8}$ fb (29.2±0.1 %)	$8.00\pm0.01\times10^{7}$ fb (81.1±0.1 %)
Jet $\lvert\eta\rvert < 1.7$	$4.38\pm0.03\times10^{7}$ fb (79.1±0.6 %)	$1.05\pm0.01\times10^{8}$ fb (76.8±0.1 %)	$6.75\pm0.01\times10^{7}$ fb (84.4±0.1 %)
Leading track $p_T > 20$ GeV/c	$3.03\pm0.03\times10^{7}$ fb (69.1±0.6 %)	$7.64\pm0.02\times10^{7}$ fb (72.7±0.2 %)	$5.40\pm0.01\times10^{7}$ fb (79.9±0.1 %)
Ch. track isolation, $p_T^{track} > 1.0$ GeV/c	$7.05\pm0.42\times10^{5}$ fb (2.3±0.1 %)	$1.94\pm0.03\times10^{6}$ fb (2.53±0.04 %)	$1.55\pm0.01\times10^{6}$ fb (2.87±0.02 %)
1 track in signal cone	$4.35\pm0.33\times10^{5}$ fb (61.8±4.6 %)	$1.14\pm0.02\times10^{6}$ fb (59.1±1.1 %)	$1.11\pm0.01\times10^{6}$ fb (71.7±0.7 %)
Leading track quality cuts	$3.94\pm0.31\times10^{5}$ fb (90.4±7.1 %)	$9.98\pm0.20\times10^{5}$ fb (87.2±1.7 %)	$9.44\pm0.10\times10^{5}$ fb (84.9±0.9 %)
Em. isolation $E_T^{ECAL} < 1.8$ GeV	$4.65\pm1.07\times10^{4}$ fb (11.8±2.7 %)	$6.48\pm0.51\times10^{4}$ fb (6.5±0.5 %)	$2.34\pm0.16\times10^{4}$ fb (2.5±0.2 %)
Electron rejection $R_e > -0.90$	$3.18\pm0.88\times10^{4}$ fb (68.4±19.0 %)	$4.29\pm0.41\times10^{4}$ fb (66.3±6.4 %)	$1.19\pm0.11\times10^{4}$ fb (51.2±4.9 %)
$R_\tau > 0.8$	$<2.45\times10^{3}$ fb	$4.77\pm1.38\times10^{3}$ fb (11.1±3.2 %)	$1.41\pm0.39\times10^{3}$ fb (11.8±3.3 %)
τ identification efficiency			
Efficiency/event	$<7.94\times10^{-7}$	$9.65\pm2.79\times10^{-6}$	$1.40\pm0.39\times10^{-5}$

Table 7.12: Cross-sections after selections and the τ-jet identification efficiency for the different p_T bins of QCD multi-jet events in the one-prong final state. The errors represent statistical uncertainty.

	$t\bar{t}$ + 0 jets	$t\bar{t}$ + 1 jet	W+3 jets, peak
Total cross-section and MC event resolution			
Cross-section (fb) / MC event	0.44	0.45	0.82
Cross-section (fb)	6.22×10^{5}	1.77×10^{5}	5.87×10^{5}
Cross-section after selections and the efficiency for the selection step			
Jet $E_T > 119$ GeV	$8.41\pm0.02\times10^{4}$ fb (13.9±0.1 %)	$6.96\pm0.02\times10^{4}$ fb (39.6±0.1 %)	$2.22\pm0.01\times10^{4}$ fb (4.60±0.03 %)
Jet $\lvert\eta\rvert < 1.7$	$7.51\pm0.02\times10^{4}$ fb (89.2±0.2 %)	$6.00\pm0.02\times10^{4}$ fb (86.3±0.2 %)	$1.84\pm0.01\times10^{4}$ fb (82.7±0.6 %)
Leading track $p_T > 20$ GeV/c	$6.16\pm0.02\times10^{4}$ fb (82.1±0.2 %)	$4.96\pm0.01\times10^{4}$ fb (82.6±0.2 %)	$1.52\pm0.01\times10^{4}$ fb (82.4±0.6 %)
Ch. track isolation, $p_T^{track} > 1.0$ GeV/c	$7.24\pm0.06\times10^{3}$ fb (11.7±0.1 %)	$3.57\pm0.04\times10^{3}$ fb (7.2±0.1 %)	675±24 fb (4.4±0.2 %)
1 track in signal cone	$6.15\pm0.05\times10^{3}$ fb (84.9±0.7 %)	$2.92\pm0.04\times10^{3}$ fb (81.8±1.0 %)	443±19 fb (65.6±2.8 %)
Leading track quality cuts	$5.72\pm0.05\times10^{3}$ fb (93.1±0.8 %)	$2.66\pm0.03\times10^{3}$ fb (91.0±1.2 %)	385±18 fb (87.0±4.0 %)
Em. isolation $E_T^{ECAL} < 1.8$ GeV	$3.67\pm0.04\times10^{3}$ fb (64.0±0.7 %)	$1.36\pm0.02\times10^{3}$ fb (51.1±0.9 %)	163±12 fb (42.3±3.0 %)
Electron rejection $R_e > -0.90$	350±12 fb (9.5±0.3 %)	163±9 fb (12.0±0.6 %)	10.7±3.0 fb (6.5±1.8 %)
$R_\tau > 0.8$	62.4±5.3 fb (17.8±1.5 %)	27.2±3.5 fb (16.8±2.1 %)	1.64±1.16 fb (15.4±10.9 %)
τ-jet identification efficiency and τ-jet contamination from $W^{\pm} \to \tau^{\pm}\nu_{\tau}$ decays			
Efficiency/event	$1.00\pm0.08\times10^{-4}$	$1.54\pm0.20\times10^{-4}$	$2.79\pm1.97\times10^{-6}$
τ's from $W^{\pm} \to \tau^{\pm}\nu_{\tau}$	65.2±6.8 %	70.5±10.7 %	<33 %

Table 7.13: Cross-sections after selections and the τ identification efficiency for the $t\bar{t}$ + 0 jets, $t\bar{t}$ + 1 jet and on-shell W+3 jets events in the one-prong final state. The errors represent statistical uncertainty.

	W+4 jets, peak	W+3 jets, tail	W+4 jets, tail
Total cross-section and MC event resolution			
Cross-section (fb) / MC event	0.25	0.033	4.5×10^{-3}
Cross-section (fb)	1.24×10^{5}	1.06×10^{3}	200
Cross-section after selections and the efficiency for the selection step			
Jet $E_T > 119$ GeV	$1.15\pm0.01\times10^{4}$ fb (10.3±0.1 %)	101±2 fb (10.9±0.2 %)	30.3±0.4 fb (16.3±0.2 %)
Jet $\|\eta\| < 1.7$	$9.55\pm0.05\times10^{3}$ fb (83.1±0.4 %)	83.5±1.7 fb (82.4±1.6 %)	25.1±0.3 fb (82.9±1.1 %)
Leading track $p_T > 20$ GeV/c	$7.96\pm0.05\times10^{3}$ fb (83.4±0.5 %)	76.0±1.6 fb (91.0±1.9 %)	22.2±0.3 fb (88.4±1.3 %)
Ch. track isolation, $p_T^{track} > 1.0$ GeV/c	307±9 fb (3.9±0.1 %)	31.4±1.0 fb (41.3±1.3 %)	5.73±0.16 fb (25.9±0.7 %)
1 track in signal cone	188±7 fb (61.3±2.3 %)	29.3±1.0 fb (93.4±3.1 %)	5.27±0.15 fb (92.0±2.7 %)
Leading track quality cuts	165±6 fb (87.6±3.4 %)	27.9±1.0 fb (95.3±3.3 %)	4.93±0.15 fb (93.4±2.8 %)
Em. isolation $E_T^{ECAL} < 1.8$ GeV	41.8±3.3 fb (25.3±2.0 %)	20.4±0.8 fb (72.9±2.9 %)	3.66±0.13 fb (74.3±2.6 %)
Electron rejection $R_e > -0.90$	5.09±1.14 fb (12.2±2.7 %)	1.44±0.22 fb (7.1±1.1 %)	0.35±0.04 fb (9.6±1.1 %)
$R_\tau > 0.8$	0.76±0.44 fb (15.0±8.7 %)	0.39±0.11 fb (27.3±7.9 %)	$7.19\pm1.80\times10^{-2}$ fb (20.5±5.1 %)
τ-jet identification efficiency and τ-jet contamination from $W^{\pm} \to \tau^{\pm}\nu_{\tau}$ decays			
Efficiency/event	$6.16\pm3.56\times10^{-6}$	$3.72\pm1.07\times10^{-4}$	$3.60\pm0.90\times10^{-4}$
τ's from $W^{\pm} \to \tau^{\pm}\nu_{\tau}$	<25 %	58.3±22.0 %	75.0±21.7 %

Table 7.14: Cross-sections after selections and the τ-jet identification efficiency for the off-shell W+3 jets and W+4 jets events in the one-prong final state. The errors represent statistical uncertainty.

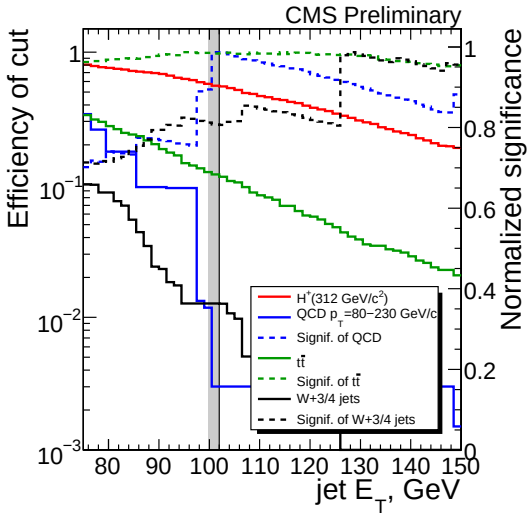

Figure 7.12: Selection efficiency (solid curves) and normalized significance (dashed curves) for three-prong τ jet candidates as a function of jet $\mathrm{E_T}$ after all other selections for finding a low bound.

τ decay modes without π^0's, the charged tracks acquire a greater fraction of the $\mathrm{a_1}$ energy than in the decay modes, where π^0's are present. Thus, the charged tracks are more collimated for decays without π^0's and a small signal cone can be selected. The maximum radius $\Delta\mathrm{R}$ around the leading track direction needed to contain all three charged tracks of signal τ decays is shown in the left plot of Fig. 7.13 after the $\mathrm{E_T}$ threshold and neutral particle rejection have been applied. It can be seen in the figure, that a signal cone size of $\mathrm{R_s} = 0.04$ will house the three tracks of the τ decay independent of the $\mathrm{m_{H^{\pm}}}$ value. Smaller signal cone values down to $\mathrm{R_s} = 0.02$ were also studied, but they were not found to improve significantly the rejection of background events.

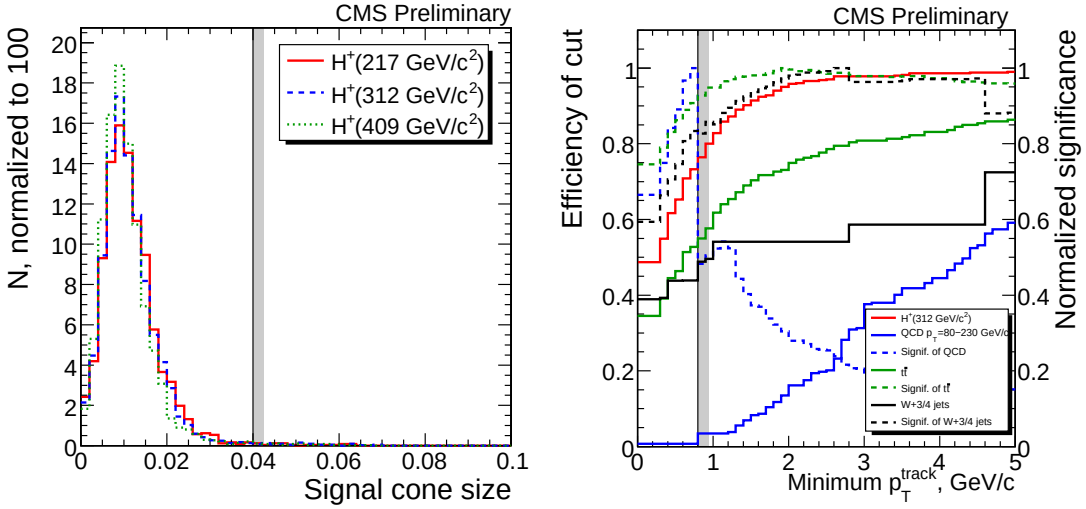

Figure 7.13: Left: The maximum distance of the tracks from the leading track direction in $\Delta\mathrm{R}$, i.e. the signal cone size containing the three charged tracks of MC matched three-prong τ decays for the signal samples after the $\mathrm{E_T} > 102$ GeV threshold and the neutral particle rejection. Right: Selection efficiency (solid curves) and normalized significance (dashed curves) for three-prong τ-jet candidates as a function of minimum track $\mathrm{p_T}$in isolation cone after all other selections for finding a low bound.

After requiring three tracks in the signal cone, no tracks were allowed between the signal cone and the outer edge of the isolation cone. The minimum transverse momentum, $\mathrm{p_T^{min}}$, of the tracks, which were counted, was varied to obtain the efficiencies of the charged track algorithm. The optimal charged track isolation performance was found to be give with $\mathrm{p_T^{min}}$ = 0.8 GeV/c (right plot of Fig. 7.13).

In addition to the track quality cuts, which were used for the one-prong τ-jet candidates, the sum of charge of the three tracks was required to be $\pm$1.

7.7.1.5.3 Electromagnetic isolation

The electromagnetic isolation described in Section 5.1.5 was applied also for the three-prong final states. It was found, that the signal cone size of 0.10, which was chosen for the one-prong τ-jet candidates, was no longer enough to accommodate the showering of the three charged pions, which can begin in the ECAL. Thus, the signal cone size was increased to 0.15 for the three-prong τ-jet candidates. The outer bound of the isolation annulus was kept at 0.50, which is the same as for the one-prong τ-jet candidates.

The electromagnetic isolation efficiency and the corresponding significance curves for the different event samples are shown for the three-prong τ-jet candidates in Fig. 7.14 as a function of $\mathrm{P_{isol}}$. It was observed, that placing a cut $\mathrm{P_{isol}} < 1.5$ GeV yielded the optimal separation against the QCD multi-jet background. This cut value was found to give a similar signal efficiency for the three-prong τ-jet candidates as for the one-prong τ-jet candidates.

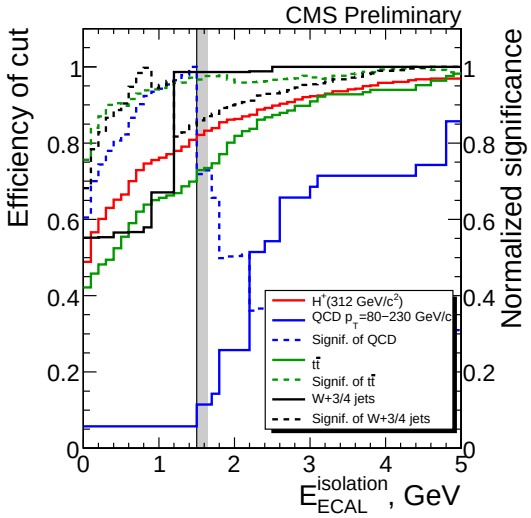

Figure 7.14: Selection efficiency (solid curves) and normalized significance (dashed curves) for three-prong τ-jet candidates as a function of energy in electromagnetic isolation annulus after all other selections for finding a low bound.

7.7.1.5.4 Flight path reconstruction

The flight path of the τ jets was reconstructed for the three-prong τ-jet candidates with the method described in Section 5.1.8. The unsigned three-dimensional flight path significance is shown in the left plot of Fig. 7.15 for three-prong τ-jet candidates from the signal and QCD multi-jet samples. In the plots, the full error matrix was taken into account. The requirement for the secondary vertex to be reconstructed in front of the primary vertex compared to the jet direction was found to reject too much signal. Majority of the b and c jets, which could fake the τ lepton flight path, were rejected by the charged track isolation as a result of the high charged track multiplicity in those jets.

The selection efficiency and the corresponding significance curves for the three-prong τ-jet candidates are shown in the left plot of Fig. 7.15 for the different event samples as a function of τ-jet candidate flight path significance. Requiring the flight path significance to exceed 2.4 was observed to give the optimal performance against QCD multi-jet events. In addition to rejecting hadronic jets from QCD multi-jet events, the flight path selection is also found to reject much of the hadronic jet content of the $\mathrm{t\bar{t}}$ background. This caused the $\mathrm{t\bar{t}}$ events, which survived the flight path selection, to consist almost solely of τ jets from the $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decay.

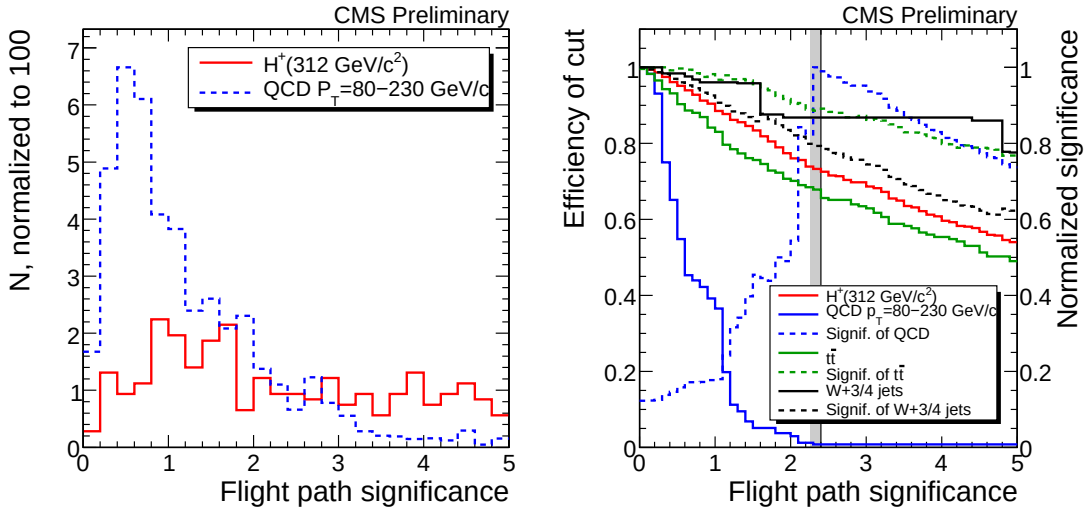

Figure 7.15: Left: The reconstructed τ flight path significance for the MC matched τ jets with $\mathrm{m}_{\mathrm{H}^{\pm}} = 312$ GeV/c^2 (solid) and hadronic jets from QCD multi-jet events with $\hat{\mathrm{p}}_{\mathrm{T}} = 80$-$230$ GeV/c (dashed) after all selections from jet E_{T} threshold to electromagnetic isolation and the neutral particle rejection. Right: Selection efficiency (solid curves) and normalized significance (dashed curves) for three-prong τ jet candidates as a function of flight path significance after all other selections for finding a low bound.

7.7.1.5.5 Invariant mass reconstruction

The τ-jet invariant mass was reconstruction has been described in Section 5.1.9. The visible τ-jet mass was calculated from the invariant mass of the charged tracks.The distribution of the invariant mass of the τ jets is shown in the left plot of Fig. 7.16 for the three-prong τ-jet candidates from signal and

QCD multi-jet samples. It was observed (right plot of Fig. 7.16), that the optimal rejection against QCD multi-jet events was achieved by requiring m_{τ} <1.5 GeV/c^2.

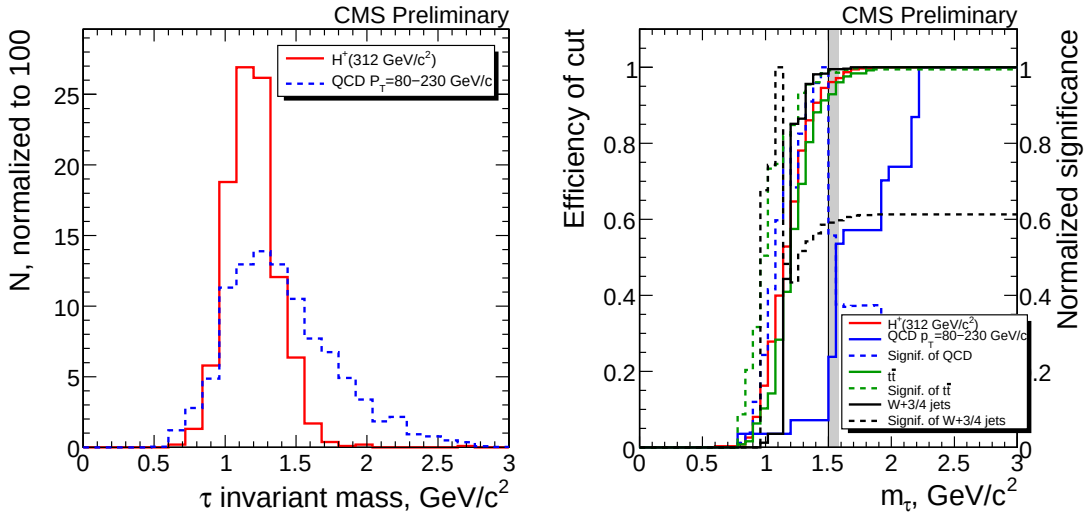

Figure 7.16: Left: The invariant visible τ-jet mass for MC matched τ jets with $m_{H^{\pm}}$ = 312 GeV/c^2 (solid) and hadronic jets from QCD multi-jet events with $\hat{p}_T$ = 80-230 GeV/c (dashed) after all selections from jet E_T threshold to electromagnetic isolation and the neutral particle rejection. Right: Selection efficiency (solid curves) and normalized significance (dashed curves) for three-prong τ-jet candidates as a function of invariant τ mass after all other selections for finding an upper bound.

7.7.1.5.6 Neutral particle rejection

Since the strategy for the three-prong final state is to select the τ decay modes without π^0's, the number of neutral particles is expected to be small in the signal τ jets and the energy carried by the charged tracks should match well the visible τ-jet energy. Hence, matching of the energy carried by the tracks to the visible τ-jet energy was required as described in Section 5.1.11. The distribution of the matching variable ΔE is shown in Fig. 7.17 for MC matched three-prong τ-jets coming from the $H^{\pm} \to \tau^{\pm}\nu_{\tau}$ decays with and without π^0's. Since the peak of the ΔE distribution in the $\tau^{\pm} \to \pi^{\pm}\pi^{\pm}\pi^{\mp}\nu_{\tau}$ decay mode was observed to be shifted to positive values, an asymmetric cut was chosen for ΔE instead of a symmetric one around ΔE = 0. To retain as much signal as possible, a window of $\Delta E >$ -0.2 was selected.

The selection efficiency and the corresponding significance curves for the neutral particle rejection are shown in the right plot Fig. 7.17 as a function of ΔE. It can be seen, that the selection $\Delta E >$ -0.2 is optimal against the QCD multi-jet background. It should be noted, that the R_{τ} variable provides additional rejection against decay modes with π^0's.

7.7.1.5.7 Helicity correlations

The τ helicity correlations have been described for the three prong final states in Sections 2.3.2.2 and 5.1.12. The selection efficiency for the τ helicity correlations and the corresponding significance

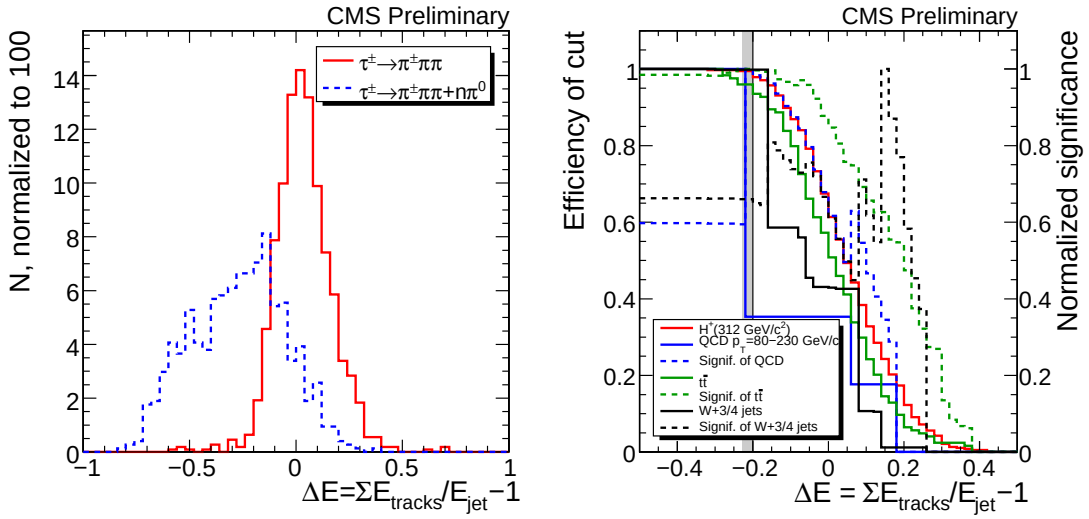

Figure 7.17: Left: The ΔE distribution for the $\tau^{\pm} \to \pi^{\pm}\pi^{\pm}\pi^{\mp}$ (solid line) and $\tau^{\pm} \to \pi^{\pm}\pi^{\pm}\pi^{\mp} + n\pi^{0}$ (dashed line) MC matched decay modes for τ jets with $m_{H^{\pm}}$ = 312 GeV/c^2 after all selections from jet E_T threshold to electromagnetic isolation. Right: Selection efficiency (solid curves) and normalized significance (dashed curves) for three-prong τ-jet candidates as a function of $\Delta E = \sum E_{tracks}/E_{jet}$ - 1 after all other selections for finding a low bound.

curves are shown in Fig. 7.18 for the three-prong τ-jet candidates from the different event samples as a function of R_{τ}. Since the remaining events of the $t\bar{t}$ and W+jets samples consist at this point almost solely of τ jets coming from the $W^{\pm} \to \tau^{\pm}\nu_{\tau}$ decay, the difference between the differently polarized τ's is clearly visible in the Fig. 7.18 as a smaller efficiency for τ's from the $W^{\pm} \to \tau^{\pm}\nu_{\tau}$ decay than for the signal τ's. The optimal cut value both against hadronic jets from QCD multi-jet background and τ jets from the $t\bar{t}$ background was observed to be $R_{\tau} > 0.55$.

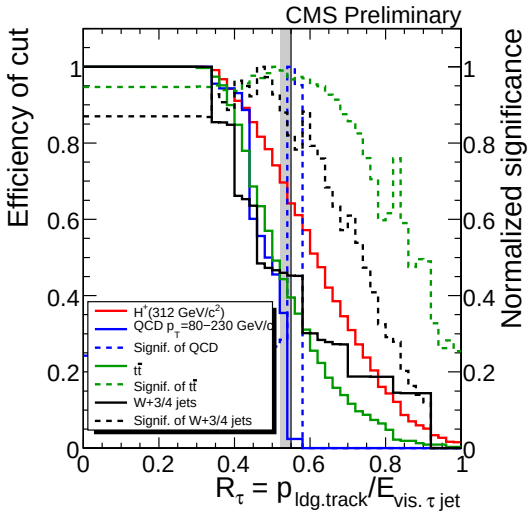

Figure 7.18: Selection efficiency (solid curves) and normalized significance (dashed curves) for three-prong τ-jet candidates as a function of R_{τ} after all other selections for finding a low bound.

7.7.1.5.8 Results for three-prong τ identification

The results of the selection criteria covered in the preceding sections for the three-prong τ selection are summarized in the Tables 7.15-7.18.

For signal events, the cross-section after each selection and the total τ selection efficiency are shown in Table 7.15. The overall selection efficiency is observed to range from $\sim$0.4% to $\sim$1.0% with efficiency rising as a function of $\mathrm{m_{H^{\pm}}}$. The results correspond to an increase of $\sim$20% of signal events when the three-prong final state is used in addition to the one-prong final state. The purity of the signal τ's is found to be 99.0% or better for the three-prong final state with no τ's from the associated $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decay passing the τ-jet identification for the higher $\mathrm{m_{H^{\pm}}}$ values.

Table 7.16 shows the cross-section after each selection and the total τ selection efficiency for QCD multi-jet events. The $\mathrm{E_T}$ threshold, charged track isolation, and the R_{τ} selections were observed to be the most effective cuts. In addition, the flight path and the τ invariant mass selections were found to yield a combined rejection factor of $\sim$10. The overall rejection factor was found to be better than 10^6 against QCD multi-jet events with $\hat{\mathrm{p}}_{\mathrm{T}}$ = 80-170 GeV/c and $\sim 5\times 10^5$ for $\hat{\mathrm{p}}_{\mathrm{T}}$ = 170-230 GeV/c.

The cross-sections after each selection and the total τ selection efficiency for $\mathrm{t\bar{t}}$ events are shown in the first two columns of Table 7.17. The most effective selections were observed to be also in this case the $\mathrm{E_T}$ threshold, charged track isolation, and the R_{τ} cut. The overall rejection factor was found to be $\sim 1.5\times 10^4$. Since the rejection against hadronic jets is very good, only $\sim$10% of the selected τ jets in the $\mathrm{t\bar{t}}$ events surviving the τ identification consist of a hadronic jet and the rest consist of τ's from the $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decay.

The last column of Table 7.17 and the Table 7.18 show the cross-sections after each selection and the τ selection efficiency for the W+jets backgrounds. The rejection factor is observed to be $\sim 4\times 10^5$ and $\sim 2\times 10^5$ for the on-shell W+3 jets and W+4 jets backgrounds, respectively. The τ jets coming from the $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decay are harder in the off-shell W + jets samples than in the on-shell ones. Therefore, the selections were less effective for the off-shell $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decays than for the on-shell ones and the overall rejection factor of three-prong τ identification is found to be $\sim 5\times 10^3$ and $\sim 9\times 10^3$ for the off-shell W+3 jets and W+4 jets backgrounds, respectively. Similarly as for the $\mathrm{t\bar{t}}$ background, the efficient rejection of hadronic jets causes the majority of the W+jets events surviving the three-prong τ identification to consist of τ's from the $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decay. Therefore, the W+jets events surviving the three-prong τ-jet identification are very signal like, especially in the off-shell $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decays.

7.7.1.6 Summary of the reoptimized τ-jet identification

The τ-jet identification of the $\mathrm{gg} \rightarrow \mathrm{tbH^{\pm}}$, $\mathrm{H}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decay channel described in Sections 7.1-7.6 and included in the physics TDRs was redone with newer software, more statistics, and an optimization of the τ-jet identification, which takes into account the correlation between the selections. In the new study,

	$H^{\pm} \to \tau^{\pm}\nu_{\tau}$ $m_{H^{\pm}}$=217 GeV/c^2 m_A=200 GeV/c^2	$H^{\pm} \to \tau^{\pm}\nu_{\tau}$ $m_{H^{\pm}}$=312 GeV/c^2 m_A=300 GeV/c^2	$H^{\pm} \to \tau^{\pm}\nu_{\tau}$ $m_{H^{\pm}}$=409 GeV/c^2 m_A=400 GeV/c^2
Total cross-section and MC event resolution			
Cross-section (fb) / MC event	0.014	2.16×10^{-3}	7.35×10^{-4}
Cross-section (fb)	1.19×10^{3}	181	57.9
Hadronic τ jet (fb)	772	117	37.5
Cross-section after selections and the efficiency for the selection step			
Jet $E_T > 102$ GeV	381±2 fb (50.4±0.3 %)	73.4±0.4 fb (63.8±0.3 %)	27.0±0.1 fb (73.1±0.4 %)
Jet $\|\eta\| < 1.8$	352±2 fb (92.5±0.6 %)	68.7±0.4 fb (93.6±0.5 %)	25.5±0.1 fb (94.6±0.5 %)
Leading track $p_T > 20$ GeV/c	298±2 fb (84.5±0.6 %)	59.8±0.4 fb (87.0±0.5 %)	22.7±0.1 fb (88.7±0.5 %)
Ch. track isolation, $p_T^{track} > 0.8$ GeV/c	84.2±1.1 fb (28.3±0.4 %)	21.0±0.2 fb (35.2±0.4 %)	8.89±0.08 fb (39.2±0.4 %)
3 tracks in signal cone	15.4±0.5 fb (18.3±0.6 %)	4.03±0.09 fb (19.2±0.4 %)	1.70±0.04 fb (19.1±0.4 %)
Track quality cuts	13.4±0.4 fb (86.6±2.8 %)	3.49±0.09 fb (86.5±2.2 %)	1.45±0.03 fb (85.3±1.9 %)
Em. isolation, $E_T^{isol.} < 1.5$ GeV	10.1±0.4 fb (76.0±2.8 %)	2.63±0.08 fb (75.3±2.2 %)	1.07±0.03 fb (73.8±2.0 %)
$\sigma_{\text{flight path}} > 2.4$	7.48±0.32 fb (73.7±3.2 %)	1.92±0.06 fb (72.9±2.5 %)	0.78±0.02 fb (72.8±2.3 %)
$m_\tau < 1.5$ GeV/c^2	7.03±0.31 fb (94.1±4.2 %)	1.85±0.06 fb (96.3±3.3 %)	0.75±0.02 fb (95.8±3.0 %)
$\Delta E = \sum E_{tracks}/E_{jet} - 1 > -0.2$	5.02±0.26 fb (71.3±3.8 %)	1.31±0.05 fb (71.2±2.9 %)	0.58±0.02 fb (78.0±2.8 %)
$R_\tau > 0.55$	2.88±0.20 fb (57.3±4.0 %)	0.82±0.04 fb (62.6±3.2 %)	0.37±0.02 fb (63.1±2.8 %)
τ identification efficiency and purity			
Efficiency/event	$3.73\pm0.26\times10^{-3}$	$7.02\pm0.36\times10^{-3}$	$9.79\pm0.44\times10^{-3}$
Signal τ purity	99.0±0.7 %	99.7±0.3 %	99.8±0.2 %
τ's from associated $W^{\pm} \to \tau^{\pm}\nu_{\tau}$	1.0±0.7 %	<0.3 %	<0.2 %

Table 7.15: Cross-sections after selections and the τ-jet identification efficiency for the signal events with $m_{H^{\pm}}$ = 217-409 GeV/c^2 in the three-prong final state. The errors represent statistical uncertainty.

	QCD $\hat{p}_T$= 80-120 GeV/c	QCD $\hat{p}_T$= 120-170 GeV/c	QCD $\hat{p}_T$= 170-230 GeV/c
Total cross-section and MC event resolution			
Cross-section (fb) / MC event	2.45×10^{3}	397	109
Cross-section (fb)	3.08×10^{9}	4.94×10^{8}	1.01×10^{8}
Cross-section after selections and the efficiency for the selection step			
Jet $E_T > 102$ GeV	$1.77\pm0.01\times10^{8}$ fb (6.52±0.02 %)	$2.46\pm0.01\times10^{8}$ fb (52.6±0.1 %)	$8.83\pm0.01\times10^{7}$ fb (89.5±0.1 %)
Jet $\lvert\eta\rvert < 1.8$	$1.41\pm0.01\times10^{8}$ fb (79.8±0.3 %)	$2.01\pm0.01\times10^{8}$ fb (81.5±0.1 %)	$7.87\pm0.01\times10^{7}$ fb (89.1±0.1 %)
Leading track $p_T > 20$ GeV/c	$9.01\pm0.05\times10^{7}$ fb (63.9±0.3 %)	$1.39\pm0.01\times10^{8}$ fb (69.2±0.1 %)	$6.24\pm0.01\times10^{7}$ fb (79.3±0.1 %)
Ch. track isolation, $p_T^{track} > 0.8$ GeV/c	$1.79\pm0.07\times10^{6}$ fb (2.0±0.1 %)	$2.90\pm0.03\times10^{6}$ fb (2.09±0.02 %)	$1.67\pm0.01\times10^{6}$ fb (2.68±0.02 %)
3 tracks in signal cone	$2.62\pm0.25\times10^{5}$ fb (14.6±1.4 %)	$3.43\pm0.12\times10^{5}$ fb (11.8±0.4 %)	$1.23\pm0.04\times10^{5}$ fb (7.4±0.2 %)
Track quality cuts	$2.40\pm0.24\times10^{5}$ fb ($91.6^{+8.4}_{-9.3}$ %)	$2.89\pm0.11\times10^{5}$ fb (84.4±3.1 %)	$1.04\pm0.03\times10^{5}$ fb (84.4±2.7 %)
Em. isolation, $E_T^{isol.} < 1.5$ GeV	$1.10\pm0.16\times10^{5}$ fb (45.9±6.8 %)	$1.03\pm0.06\times10^{5}$ fb (35.6±2.2 %)	$3.49\pm0.19\times10^{4}$ fb (33.6±1.9 %)
$\sigma_{\text{flight path}} > 2.4$	$2.20\pm0.73\times10^{4}$ fb (20.0±6.7 %)	$3.06\pm0.35\times10^{4}$ fb (29.7±3.4 %)	$9.78\pm1.03\times10^{3}$ fb (28.0±3.0 %)
$m_\tau < 1.5$ GeV/c^2	$1.96\pm0.69\times10^{4}$ fb ($88.9^{+11.1}_{-31.4}$ %)	$2.42\pm0.31\times10^{4}$ fb (79.2±10.1 %)	$7.17\pm0.88\times10^{3}$ fb (73.3±9.0 %)
$\Delta E = \sum E_{tracks}/E_{jet} - 1 > -0.2$	$4.89\pm3.46\times10^{3}$ fb (25.0±17.7 %)	$3.18\pm1.12\times10^{3}$ fb (13.1±4.6 %)	869±307 fb (12.1±4.3 %)
$R_\tau > 0.55$	$<2.45\times10^{3}$ fb	<397 fb	217±154 fb (25.0±17.7 %)
τ identification efficiency			
Efficiency/event	$<7.94\times10^{-7}$	$<8.04\times10^{-7}$	$2.15\pm1.52\times10^{-6}$

Table 7.16: Cross-sections after selections and the τ-jet identification efficiency for the different p_T bins of QCD multi-jet events in the three-prong final state. The errors represent statistical uncertainty.

	$t\bar{t}$ + 0 jets	$t\bar{t}$ + 1 jet	W+3 jets, peak
Total cross-section and MC event resolution			
Cross-section (fb) / MC event	0.44	0.45	0.82
Cross-section (fb)	6.22×10^{5}	1.77×10^{5}	5.87×10^{5}
Cross-section after selections and the efficiency for the selection step			
Jet $E_T > 102$ GeV	$1.34\pm0.01\times10^{5}$ fb (22.2±0.1 %)	$9.08\pm0.02\times10^{4}$ fb (51.6±0.1 %)	$3.67\pm0.02\times10^{4}$ fb (7.58±0.04 %)
Jet $\lvert\eta\rvert < 1.8$	$1.22\pm0.01\times10^{5}$ fb (91.0±0.2 %)	$8.11\pm0.02\times10^{4}$ fb (89.4±0.2 %)	$3.11\pm0.02\times10^{4}$ fb (84.9±0.4 %)
Leading track $p_T > 20$ GeV/c	$9.64\pm0.02\times10^{4}$ fb (78.9±0.2 %)	$6.54\pm0.02\times10^{4}$ fb (80.6±0.2 %)	$2.47\pm0.01\times10^{4}$ fb (79.2±0.5 %)
Ch. track isolation, $p_T^{track} > 0.8$ GeV/c	$1.06\pm0.01\times10^{4}$ fb (11.0±0.1 %)	$4.62\pm0.05\times10^{3}$ fb (7.1±0.1 %)	$1.76\pm0.04\times10^{3}$ fb (7.1±0.2 %)
3 tracks in signal cone	629±17 fb (5.9±0.2 %)	281±11 fb (6.1±0.2 %)	107±9 fb (6.1±0.5 %)
Track quality cuts	537±15 fb (85.3±2.4 %)	236±10 fb (83.9±3.7 %)	92.6±8.7 fb (86.3±8.1 %)
Em. isolation, $E_T^{isol.} < 1.5$ GeV	292±11 fb (54.4±2.1 %)	118±7 fb (50.0±3.1 %)	36.1±5.4 fb (38.9±5.9 %)
$\sigma_{\text{flight path}} > 2.4$	174±9 fb (59.5±3.0 %)	68.8±5.5 fb (58.3±4.7 %)	9.84±2.84 fb (27.3±7.9 %)
$m_\tau < 1.5$ GeV/c^2	162±8 fb (93.1±4.9 %)	62.5±5.3 fb (90.9±7.7 %)	9.02±2.72 fb ($91.7^{+8.3}_{-27.6}$ %)
$\Delta E = \sum E_{\text{tracks}}/E_{\text{jet}} - 1 > -0.2$	97.8±6.6 fb (60.4±4.1 %)	40.6±4.3 fb (65.0±6.8 %)	4.10±1.83 fb (45.5±20.3 %)
$R_\tau > 0.55$	37.6±4.1 fb (38.5±4.2 %)	13.8±2.5 fb (34.1±6.1 %)	1.64±1.16 fb (40.0±28.3 %)
τ-jet identification efficiency and τ-jet contamination from $W^{\pm} \to \tau^{\pm}\nu_{\tau}$ decays			
Efficiency/event	$6.05\pm0.66\times10^{-5}$	$7.81\pm1.40\times10^{-5}$	$2.79\pm1.97\times10^{-6}$
τ's from $W^{\pm} \to \tau^{\pm}\nu_{\tau}$	$95.3^{+4.7}_{-10.6}$ %	$90.3^{+9.7}_{-17.1}$ %	50.0±50.0 %

Table 7.17: Cross-sections after selections and the τ-jet identification efficiency for the $t\bar{t}$ +0 jets, $t\bar{t}$ + 1 jet and on-shell W+3 jets events in the three-prong final state. The errors represent statistical uncertainty.

	W+4 jets, peak	W+3 jets, tail	W+4 jets, tail
Total cross-section and MC event resolution			
Cross-section (fb) / MC event	0.25	0.033	4.5×10^{-3}
Cross-section (fb)	1.24×10^{5}	1.06×10^{3}	200
Cross-section after selections and the efficiency for the selection step			
Jet $\mathrm{E_T} > 102$ GeV	$1.69\pm0.01\times10^{4}$ (15.21 ± 0.06 %)	159 ± 2 (17.2 ± 0.2 %)	44.6 ± 0.4 (24.1 ± 0.2 %)
Jet $\lvert\eta\rvert < 1.8$	$1.44\pm0.01\times10^{4}$ (85.6 ± 0.4 %)	134 ± 2 (84.4 ± 1.3 %)	38.0 ± 0.4 (85.2 ± 0.9 %)
Leading track $\mathrm{p_T} > 20$ GeV/c	$1.16\pm0.01\times10^{4}$ (80.1 ± 0.4 %)	120 ± 2 (89.4 ± 1.5 %)	32.7 ± 0.4 (86.2 ± 1.0 %)
Ch. track isolation, $\mathrm{p_T^{track}} > 0.8$ GeV/c	602 ± 12 (5.2 ± 0.1 %)	45.6 ± 1.2 (38.0 ± 1.0 %)	8.53 ± 0.20 (26.0 ± 0.6 %)
3 tracks in signal cone	41.3 ± 3.2 (6.8 ± 0.5 %)	1.71 ± 0.24 (3.7 ± 0.5 %)	0.37 ± 0.04 (4.3 ± 0.5 %)
Track quality cuts	33.9 ± 2.9 (82.1 ± 7.1 %)	1.51 ± 0.22 ($88.5^{+11.5}_{-13.0}$ %)	0.32 ± 0.04 (87.8 ± 10.3 %)
Em. isolation, $\mathrm{E_T^{isol.}} < 1.5$ GeV	14.0 ± 1.9 (41.4 ± 5.6 %)	1.21 ± 0.20 (80.4 ± 13.2 %)	0.24 ± 0.03 (75.0 ± 10.2 %)
$\sigma_{\mathrm{flight\ path}} > 2.4$	5.09 ± 1.14 (36.4 ± 8.1 %)	0.82 ± 0.16 (67.6 ± 13.5 %)	0.18 ± 0.03 (72.2 ± 11.6 %)
$\mathrm{m}_{\tau} < 1.5$ GeV/c^2	4.33 ± 1.05 ($85.0^{+15.0}_{-20.6}$ %)	0.72 ± 0.15 ($88.0^{+12.0}_{-18.8}$ %)	0.15 ± 0.03 ($87.2^{+12.8}_{-15.0}$ %)
$\Delta\mathrm{E} = \sum \mathrm{E_{tracks}}/\mathrm{E_{jet}} - 1 > -0.2$	1.02 ± 0.51 (23.5 ± 11.8 %)	0.56 ± 0.14 (77.3 ± 18.7 %)	0.10 ± 0.02 (67.6 ± 14.1 %)
$\mathrm{R}_{\tau} > 0.55$	0.76 ± 0.44 ($75.0^{+25.0}_{-43.3}$ %)	0.20 ± 0.08 (35.3 ± 14.4 %)	$2.25\pm1.01\times10^{-2}$ (21.7 ± 9.7 %)
τ-jet identification efficiency and τ-jet contamination from $\mathrm{W}^{\pm} \to \tau^{\pm}\nu_{\tau}$ decays			
Efficiency/event	$6.16\pm3.56\times10^{-6}$	$1.86\pm0.76\times10^{-4}$	$1.12\pm0.50\times10^{-4}$
τ's from $\mathrm{W}^{\pm} \to \tau^{\pm}\nu_{\tau}$	$66.7^{+33.3}_{-47.1}$ %	100^{+0}_{-41} %	100^{+0}_{-45} %

Table 7.18: Cross-sections after selections and the τ-jet identification efficiency for the off-shell W+3 jets and W+4 jets events in the three-prong final state. The errors represent statistical uncertainty.

the jet $\mathrm{E_T}$ cut was set to a higher value than in the physics TDR study, the single τ trigger without the cut on $\mathrm{E_T^{miss}}$ was included in the τ-jet identification efficiencies, and the cut on $\mathrm{E_T^{miss}}$ was not applied. Furthermore, τ-jet identification selections based on reconstruction of the τ-lepton flight path and the visible τ-lepton mass were included in the new study.

Since the cut on the $\mathrm{E_T^{miss}}$ can be assumed to be relatively uncorrelated with the τ-jet identification, a rough comparison between the physics TDR study and the reoptimization study can be made. Hence, if both the one- and three-prong final states are considered, and if the slight difference in the $\mathrm{m_{H^{\pm}}}$ mass points between the studies is taken into account, the efficiency of the reoptimized τ-jet identification is found to be similar for the signal to that obtained in Section 7.3.4, with a slight improvement on the efficiency of the $\mathrm{m_{H^{\pm}}}$ = 217 GeV/c^2 signal sample. The τ-jet identification efficiency obtained for the $\mathrm{t\bar{t}}$ background with the reoptimized study can be estimated to remain almost the same as in the physics TDR study. On the other hand, the W+ 3 jets and the QCD multi-jet events are found to be further suppressed by factors of $\sim$40 and $\sim$14, respectively, compared to the physics TDR study. These results obtained with the reoptimized study thus confirm the outcome of the physics TDR study: it is possible to suppress the QCD multi-jet background enough to extract the mass peak for the heavy charged MSSM Higgs boson.

One of the open questions with the τ-jet identification in the $\mathrm{H}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ channel has been whether or not the three-prong final state can be used to increase the significance of the signal. The inclusion of the three-prong final state has been questioned because of the large hadronic jet background from the QCD multi-jet events. This study showed, that tight cuts can suppress the QCD multi-jet events enough for the three-prong final state to be usable. However, it was found out that the number of $\mathrm{t\bar{t}}$ events with $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decays is enhanced if the three-prong final state is used in addition to the one-prong final state. Hence, it remains to be seen in a future full study of the $\mathrm{H}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decay channel whether the $\mathrm{t\bar{t}}$ can be suppressed enough with event selections other than the τ-jet identification for the three-prong final state to be usable.

8 Conclusions

The breaking of the electroweak symmetry is the last unconfirmed part of the standard model of particle physics, which describes the known particles and the electromagnetic, weak, and strong interactions between them. The electroweak symmetry is assumed to be spontaneously broken, which yields masses to different particles through the Higgs mechanism. Hence, it would explain how the masses of the particles are generated. However, the Higgs mechanism requires the existence of a scalar Higgs boson, which so far has not been experimentally confirmed. Since the mass of the Higgs boson is limited as a result of theoretical constraints below the TeV energy scale, the LHC and its general-purpose detectors, including the CMS experiment, are, for the first time, in a position to either confirm or exclude the existence of the Higgs boson.

One of the flaws of the standard model, which otherwise works remarkably well, is that the self-couplings of the Higgs boson are known to cause divergences at high energies. To remedy this problem, the most prominent theory to extend the standard model to higher energies is the minimal supersymmetric standard model (MSSM), which introduces a symmetry between fermions and bosons. The price of the new symmetry is, however, that a supersymmetrical partner should exist for each known particle. Additionally, the Higgs sector has to be extended to contain two Higgs doublets which yield five physical Higgs bosons instead of just one. Since the supersymmetrical particles are assumed to be rather heavy, and since their detection is experimentally challenging, the discovery of the Higgs bosons and the measurement of their properties present an excellent opportunity to constrain the theories beyond the standard model.

Since the Higgs mechanism generates masses to particles, the Higgs fields couple to all massive particles. Hence, the Higgs bosons may decay to a plethora of final states. Of these final states, the processes $\mathrm{gg} \to \mathrm{b\bar{b}H(A)}$, $\mathrm{H, A} \to \tau\tau$ and $\mathrm{gg} \to \mathrm{tbH^{\pm}}$, $\mathrm{H^{\pm}} \to \tau^{\pm}\nu_{\tau}$, where the τ's decay hadronically, were studied in this thesis. The $\mathrm{H, A} \to \tau\tau$ decay mode has a significant branching fraction in a large part of the MSSM parameter space and it is experimentally one of the most promising decay channels for discovering the heavy neutral MSSM Higgs bosons, especially at high $\tan\beta$ values. The $\mathrm{H^{\pm}} \to \tau^{\pm}\nu_{\tau}$ is the prime channel for discovering the charged MSSM Higgs bosons.

The event selection of both the $\mathrm{H, A} \to \tau\tau$ and $\mathrm{H^{\pm}} \to \tau^{\pm}\nu_{\tau}$ decay modes relies on the identification of energetic hadronically decaying τ leptons, i.e. τ jets, the identification of associated b jets, the reconstruction of the missing $\mathrm{E_T}$, and a veto on central jets. The event selection of the $\mathrm{H^{\pm}} \to \tau^{\pm}\nu_{\tau}$ decay channel also exploits the reconstruction of the associated top quark and W masses and a veto on electrons and muons to suppress additional neutrinos in the events. Of these event selection methods, the identification of the τ jets is the single most important one.

The τ-jet identification was evaluated against the hadronic jets from QCD multi-jet events, which constitute a large but reducible background. The charged track isolation, in which the charged particle tracks are contained within a narrow signal cone around the τ-jet axis and no tracks are required to be found in an isolation annulus outside the signal cone, combined with criteria for track quality and a cut on the

number of signal tracks, was found to be the single most effective identification method for rejecting background. The second most effective identification method was found to be electromagnetic isolation, in which only minimal electromagnetic energy deposition is allowed in between the signal and the isolation cones. Strategies were presented to use a light version of the charged track isolation and electromagnetic isolation criteria to trigger on single or double τ jets in an event. For the $\mathrm{H}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decay channel, the helicity correlations were found to be important to separate signal τ jets from τ jets coming from the $\mathrm{W}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decays in electroweak backgrounds, but the algorithm was also found to enhance the suppression of hadronic jets from QCD multi-jet events. The τ-jet identification methods based on the non-zero leading track impact parameter, the reconstruction of the visible τ mass, and the rejection of electrons, muons, and neutral hadrons were studied. These identification methods are applicable to both one- and three-prong τ decay final states. Furthermore, the τ-jet identification with the τ flight path reconstruction was studied in detail for the three-prong τ decays. Despite the highly collimated tracks, the flight path based τ identification algorithm was found to be usable.

Since many of the τ-jet identification methods depend on the reconstructed tracks, the systematic uncertainties arising from the mechanical displacement of the tracking sensors were evaluated in detail. The analysis revealed that the sensor displacement increases the probability that a reconstructed track is constructed from hits belonging to different simulated tracks. This effect was found to increase the number of jets passing the charged track isolation criterion, especially in the one-prong final state. It was found out that these "fake" tracks could be strongly suppressed with track quality cuts and with an upper limit cut on the impact parameter of the leading track. The overall uncertainty resulting from the sensor displacement was found to be a few percent in the cut based on the reconstructed flight path after the charge track isolation criterion.

The τ-jet identification methods were applied together with the standard event selections to the $\mathrm{H}, \mathrm{A} \rightarrow \tau\tau$ and $\mathrm{H}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decay channels. The identification method based on the lower limit cut of the leading track impact parameter of the τ-jet candidates was not found useful, but instead the upper limit cut on the transverse impact parameter was incorporated to the track quality requirements. The signal efficiency for $\mathrm{H}, \mathrm{A} \rightarrow \tau\tau$ events to pass τ-jet identification on both τ's at was found to be 1.0-7.9% for $\mathrm{m_A}$ = 200, 500, and 800 GeV/c^2. The double τ trigger and the cut on the transverse energy of the τ jets were included in the efficiencies. The three-prong τ decay final states were only included in the study with $\mathrm{m_A}$ = 200 GeV/c^2. This selection yielded a rejection factor of 2.0-5.4$\times 10^5$ against QCD multi-jet events with $\hat{\mathrm{p}}_{\mathrm{T}} > 80$ GeV/c.

In the $\mathrm{H}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ channel, a much tighter τ-jet identification is needed to reduce the QCD multi-jet background to a tolerable level compared to the τ-jet identification in the $\mathrm{H}, \mathrm{A} \rightarrow \tau\tau$ channel. The efficiency for τ jets from the $\mathrm{H}^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decays to pass the reoptimized τ-jet identification was found to be (1.7$\pm$0.1)-(4.6$\pm$0.1)% for $\mathrm{m_A}$ =200, 300, and 400 GeV/c^2 in the one-prong τ decay final state. The cut on the transverse energy of the τ jet and the single τ trigger without the cut on the missing $\mathrm{E_T}$ have been included in the efficiencies. The rejection factors obtained with the τ-jet identification were found to be (7.1$\pm$2.0)$\times 10^4$-1.3$\times 10^6$ against the QCD multi-jet background with $80 < \hat{\mathrm{p}}_{\mathrm{T}} < 230$ GeV/c

and (1.0±0.1)×10^4-(6.5±0.8)×10^3 against the $\mathrm{t\bar{t}}$ + 0/1 jets background. The τ-jet identification of the $\mathrm{H^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}}$ analysis was also reoptimized for the three-prong τ decay final state. The signal efficiency to pass the reoptimized τ-jet identification was found to be (0.37±0.03)-(0.98±0.04)% for three-prong signal τ jets for $\mathrm{m_A}$ =200, 300, and 400 GeV/c^2. The cut on the transverse energy of the τ jet and the single τ trigger without the cut on the missing $\mathrm{E_T}$ have been included in the efficiencies. Since the three-prong τ-jet identification was observed to select more $\mathrm{t\bar{t}}$ and W+jets events compared to the signal events than in the one-prong τ-jet final state, it remains to be seen in the future full studies of the $\mathrm{H^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}}$ channel whether the three-prong final states can be used.

The τ-jet identification and the other selection methods were found to suppress the main backgrounds enough to extract the Higgs boson mass and transverse mass peaks in the $\mathrm{H, A \rightarrow \tau\tau}$ and $\mathrm{H^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}}$ decay channels, respectively. In the $\mathrm{H, A \rightarrow \tau\tau}$ decay mode, the QCD multi-jet background was found to constitute the main background and the electroweak backgrounds were found to be almost fully suppressed. In the $\mathrm{m_h^{max}}$ MSSM scenario with μ = 200 GeV/c^2, the 5σ discovery of the heavy neutral MSSM Higgs bosons was found to be possible for $\tan\beta > 20$ at $\mathrm{m_A}$ = 200 GeV/c^2 with the discovery contour rising almost linearly with $\mathrm{m_A}$ to cover $\tan\beta > 46$ at $\mathrm{m_A}$ = 800 GeV/c^2 for an integrated luminosity of 60 fb^{-1}. The estimation of systematic uncertainties as a result of the jet and missing $\mathrm{E_T}$ energy scales was found to increase the $\tan\beta$ threshold for the discovery by 6%. The $\mathrm{H, A \rightarrow \tau\tau}$ decay channel was found to be very promising for heavy neutral Higgs boson discoveries in a wide range of the permitted MSSM parameter space although it requires a relatively high integrated luminosity.

In the $\mathrm{H^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}}$ decay channel, the largest background was formed by the $\mathrm{t\bar{t}}$ events. The contribution from the QCD multi-jet and W+jets events to the total backgrounds was found to be significant, but contained. In the $\mathrm{m_h^{max}}$ MSSM scenario with μ = 200 GeV/c^2, the 5σ discovery region of the charged MSSM Higgs bosons was found to cover $\tan\beta > 29$ at $\mathrm{m_A}$ = 200 GeV/c^2 with the contour rising quadratically with $\mathrm{m_A}$ to $\tan\beta > 50$ for $\mathrm{m_A}$ = 400 GeV/c^2 for an integrated luminosity of 30 fb^{-1}. The estimation of systematic uncertainties was found to increase the $\tan\beta$ threshold for the discovery by ∼17%. The $\mathrm{H^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}}$ is hence in fact the most promising decay channel to discover the charged MSSM Higgs boson, if $\mathrm{m_{H^{\pm}} > m_t}$.

The analyses presented in this thesis rely on the assumption that the MSSM theory provides a sufficiently accurate description of the microcosm of particles and their interactions. Although the choice of the MSSM parameters has a significant effect on the cross-section of the $\mathrm{H, A \rightarrow \tau\tau}$ and $\mathrm{H^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}}$ processes, these channels provide a rather robust opportunity to discover the heavy MSSM Higgs bosons in a large part of the permitted MSSM parameter space. If the MSSM Higgs bosons exist, their discovery in the τ channels is thus a matter of collecting enough integrated luminosity.

Since the analysis and the results are based on simulations, they are prone to systematic uncertainties resulting from simplifications in the simulation, reconstruction, and analysis codes, despite extensive development and validation efforts. More recent releases of the reconstruction and analysis software are known to include a more realistic treatment of noise in the subdetectors, which has some impact on the missing $\mathrm{E_T}$ resolution. Furthermore, although the simulations were run on large data samples, some

background samples were still found to contain a limited amount of events. Hence, it was necessary to apply factorization of the relatively uncorrelated event selections to some samples, which may increase the systematic uncertainties considerably. Eventually, the systematic uncertainties will be measured with data-driven methods, of which the most important ones for the $\mathrm{H}, \mathrm{A} \to \tau\tau$ and $\mathrm{H}^{\pm} \to \tau^{\pm}\nu_{\tau}$ analyses have been sketched in this thesis.

Considerable development of the reconstruction and analysis software has taken place after the majority of the results presented in this thesis were published in the CMS physics TDRs. This development has improved the performance of many of the algorithms to reconstruct physics objects in terms of greater signal efficiency and purity. Reconstruction algorithm artefacts such as the occurrence of fake tracks discussed in this thesis have been strongly suppressed in more recent releases of the reconstruction and analysis software. Such advances are expected to benefit the $\mathrm{H}, \mathrm{A} \to \tau\tau$ and $\mathrm{H}^{\pm} \to \tau^{\pm}\nu_{\tau}$ analyses, extending the discovery reach to lower $\tan\beta$ values. Furthermore, the analysis of the $\mathrm{H}, \mathrm{A} \to \tau\tau$ and $\mathrm{H}^{\pm} \to \tau^{\pm}\nu_{\tau}$ decay channels was conducted with rectangular cuts in order to retain a transparent physical understanding of why certain selection cuts were important and others not. After the publication of the physics TDRs, many physics analyses have successfully applied multivariate techniques such as neural networks, bagged decision trees, or likelihood-based methods to improve their physics reach. It is thus expected that the $\mathrm{H}, \mathrm{A} \to \tau\tau$ and $\mathrm{H}^{\pm} \to \tau^{\pm}\nu_{\tau}$ analyses would also benefit from the multivariate techniques.

The search for the Higgs boson(s), which has been ongoing for decades, is finally drawing to a close with the advent of the LHC era in high-energy physics. Even though reality is often more complex than simulations trying to describe reality, strong arguments have been presented in this thesis that the $\mathrm{gg} \to \mathrm{b\bar{b}H(A)}$, $\mathrm{H}, \mathrm{A} \to \tau\tau$ and $\mathrm{gg} \to \mathrm{tbH}^{\pm}$, $\mathrm{H}^{\pm} \to \tau^{\pm}\nu_{\tau}$ processes can be used within reasonable uncertainties in the CMS experiment to search for the heavy neutral and charged MSSM Higgs bosons with τ-jet identification as the decisive part of the event selections.

References

[1] C. Wu *et al.*: *"Experimental test of parity conservation in beta decay"*, Phys. Rev. **105** (1957) 1413–1415.

[2] R. Feynman and M. Gell-Mann: *"Theory of the Fermi interaction"*, Phys. Rev. **109** (1958) 193–198.

[3] E. Sudarshan and R. Marshak: *"Chirality invariance and the universal Fermi interaction"*, Phys. Rev. **109** (1958) 1860–1862.

[4] Y. Nambu: *"Quasi-particles and gauge invariance in the theory of superconductivity"*, Phys. Rev. **117** (1960) 648–663.

[5] J. Goldstone, A.Salam and S. Weinberg: *"Broken Symmetries"*, Phys. Rev. **127** (1962) 965–970.

[6] P. W. Higgs: *"Broken Symmetries and the Masses of Gauge Bosons"*, Phys. Rev. Lett. **13** (1964) 508–509.

[7] F. Englert and R. Brout: *"Broken Symmetries and the Masses of Gauge Vector Mesons"*, Phys. Rev. Lett. **13** (1964) 321–323.

[8] G. S. Guralnik, C. R. Hagen and T. W. B. Kibble: *"Global Conservation Laws and Massless Particles"*, Phys. Rev. Lett. **13** (1964) 585–587.

[9] S. Weinberg: *"A Model of Leptons"*, Phys. Rev. Lett. **19** (1967) 1264–1266.

[10] S. L. Glashow: *"Partial-symmetries of weak interactions"*, Nucl. Phys. **22** (1961) 579–588.

[11] G. 't Hooft and M. Veltman: *"Regularization and renormalization of gauge fields"*, Nucl. Phys. B **44** (1972) 189–213.

[12] The UA1 Collaboration: *"Experimental observation of isolated large transverse energy electrons with associated missing energy at $\sqrt{s}$=540 GeV"*, Phys. Lett. B **122** (1983) 103–116.

[13] The UA2 Collaboration: *"Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the CERN $\bar{\mathrm{p}}\mathrm{p}$ collider"*, Phys. Lett. B **122** (1983) 476–485.

[14] The CDF Collaboration: *"Observation of Top Quark Production in $\bar{\mathrm{p}}\mathrm{p}$ Collisions with the Collider Detector at Fermilab"*, Phys. Rev. Lett. **74** (1995) 2626–2631.

[15] The DØ Collaboration: *"Observation of the Top Quark"*, Phys. Rev. Lett. **74** (1995) 2632–2637.

[16] S. Dimopoulos and H. Georgi: *"Softly Broken Supersymmetry and SU(5)"*, Nucl. Phys. B **193** (1981) 150–162.

[17] F. Halzen and A. Martin: *"Quarks & Leptons"*, (John Wiley & Sons, 1984).

[18] A. Djouadi: *"The anatomy of electroweak symmetry breaking: Tome I: The Higgs Boson in the Standard Model"*, Phys. Rep. **457** (2008) 1–216.

[19] P. W. Higgs: *"Spontaneous Symmetry Breakdown without Massless Bosons"*, Phys. Rev. **145** (1966) 1156–1163.

[20] T. Kibble: *"Symmetry Breaking in Non-Abelian Gauge Theories"*, Phys. Rev. **155** (1967) 1554–1561.

[21] The ALEPH, CDF, D0, DELPHI, L3, OPAL, SLD Collaborations, the LEP Electroweak Working Group, the Tevatron Electroweak Working Group and the SLD Electroweak and Heavy Flavour Groups: *"Precision Electroweak Measurements and Constraints on the Standard Model"*, CERN-PH-EP **020** (2008).

[22] LEP Collaborations: *"Search for the Standard Model Higgs boson at LEP"*, Phys. Lett. B **565** (2003) 61–75.

[23] B. W. Lee, C. Quigg and H. B. Thacker: *"Weak interactions at very high energies: The role of the Higgs-boson mass"*, Phys. Rev. D **16** (1977) 1519–1531.

[24] T. Hambye and K. Riesselmann: *"Matching conditions and Higgs boson mass upper bounds reexamined"*, Phys. Rev. D **55** (1997) 7255–7262.

[25] CMS Collaboration: *"CMS Physics Technical Design Report, Volume II: Physics Performance"*, J. Phys. G: Nucl. Part. Phys. **34** (2007) 995–1579.

[26] C. Kolda and H. Murayama: *"The Higgs mass and new physics scales in the minimal standard model"*, JHEP **07** (2000) 035.

[27] P. Fayet: *"Supergauge invariant extension of the Higgs mechanism and a model for the electron and its neutrino"*, Nucl. Phys. B **90** (1974) 104–124.

[28] P. Fayet: *"Relations between the masses of the superpartners of leptons and quarks, the goldstino coupling and the neutral currents"*, Phys. Lett. B **84** (1979) 416–420.

[29] K. Inoue, A. Kakuto, H. Komatsu and S. Takeshita: *"Aspects of Grand Unified Models with Softly Broken Supersymmetry"*, Prog. Theor. Phys. **68** (1982) 927–946.

[30] I. Simonsen: *"A Review of Minimal Supersymmetric Electro-Weak Theory"*, hep-ph/9506369v2 (1995).

[31] Y. Golf'and and E. Likhtman: *"Extension of the algebra of Poincaré group generators and violation of P invariance"*, JETP Lett. **13** (1971) 323–326.

[32] D. Volkov and V. Akulov: *"Is the neutrino a goldstone particle?"*, Phys. Lett. B **46** (1973) 109–110.

[33] J. Wess and B. Zumino: *"Supergauge transformations in four dimensions"*, Nucl. Phys. B **70** (1974) 39–50.

[34] U. Amaldi, W. de Boer and H. Fürstenau: *"Comparison of grand unified theories with electroweak and strong coupling constants measured at LEP"*, Phys. Lett. B **260** (1991) 447–455.

[35] G. Farrar and P. Fayet: *"Phenomenology of the production, decay, and detection of new hadronic states associated with supersymmetry"*, Phys. Lett. B **76** (1978) 575–579.

[36] A. Djouadi: *"The anatomy of electroweak symmetry breaking Tome II: The Higgs bosons in the Minimal Supersymmetric Model"*, Phys. Rep. **459** (2008) 1–241.

[37] LEP Higgs Working Group: *"Searches for the Neutral Higgs Bosons of the MSSM"*, hep-ex/0107030 (2001).

[38] J. Gunion, H. Haber, G. Kane and S. Dawson: *"The Higgs Hunter's Guide"*, (Addison-Wesley, 1990).

[39] M. S. Carena, S. Heinemeyer, C. E. M. Wagner and G. Weiglein: *"Suggestions for improved benchmark scenarios for Higgs- boson searches at LEP2"*, CERN-TH **99-374** (1999).

[40] S. Heinemeyer, W. Hollik and G. Weiglein: *"The masses of the neutral CP-even Higgs bosons in the MSSM: Accurate analysis at the two-loop level"*, Eur. Phys. J. C **9** (1999) 343–366.

[41] S. Heinemeyer and G. Weiglein: *"Higgs-mass predictions and electroweak precision observables in the Standard Model and the MSSM"*, Nucl. Phys. B **89** (2000) 216–221.

[42] S. Heinemeyer, W. Hollik and G. Weiglein: *"FeynHiggs: A program for the calculation of the masses of the neutral CP even Higgs bosons in the MSSM"*, Comp. Phys. Comm. **124** (2000) 76–89.

[43] Particle Data Group with C. Amsler *et al.*: *"Review of Particle Physics"*, Phys. Lett. B **667** (2008) 1–1340.

[44] D. P. Roy: *"The hadronic tau decay signature of a heavy charged Higgs boson at LHC"*, Phys. Lett. B **459** (1999) 607–614.

[45] D. P. Roy: *"Looking for the charged Higgs boson"*, Mod. Phys. Lett. A **19** (2004) 1813–1828.

[46] M. Guchait, R. Kinnunen and D. P. Roy: *"Signature of heavy charged Higgs boson at LHC in the 1- and 3-prong hadronic tau decay channels"*, Eur. Phys. J. C **52** (2007) 665–672.

[47] The CMS Collaboration: *"The CMS Experiment at the CERN LHC"*, JINST **3** (2008) S08004.

[48] CMS Collaboration: *"The Compact Muon Solenoid Technical Proposal"*, CERN/LHCC **1994-38** (1994).

[49] The TOTEM Collaboration: *"The TOTEM Experiment at the CERN LHC"*, JINST **3** (2008) S08007.

[50] CMS Collaboration: *"CMS: The Tracker Project Technical Design Report"*, CERN/LHCC **1998-6** (1998).

[51] CMS Collaboration: *"Addendum to the Tracker TDR by the CMS Collaboration"*, CERN/LHCC **2000-16** (2000).

[52] CMS Collaboration: *"CMS Physics Technical Design Report, Volume I: Detector Performance and software"*, CERN/LHCC **2006-001** (2006).

[53] R. Baur and W. Bertl: *"The CMS Pixel Vertex Detector"*, Nucl. Phys. B **78** (1999) 293–300.

[54] D. Kotlinski: *"The CMS pixel detector"*, Nucl. Instr. & Meth. A **465** (2001) 46–50.

[55] G. Bolla *et al.*: *"Design and test of pixel sensors for the CMS experiment"*, Nucl. Instr. & Meth. A **461** (2001) 182–184.

[56] S. Braibant *et al.*: *"Investigation of Design Parameters and Choice of Substrate Resistivity and Crystal Orientation for the CMS Silicon Microstrip Detector"*, CMS Note **2000/011** (2000).

[57] S. Abdullin *et al.*: *"Design, Performance, and Calibration of the CMS Hadron-Outer Calorimeter"*, CMS Note **2008/020** (2009).

[58] CMS Collaboration: *"CMS Electromagnetic Calorimeter Project Technical Design Raport"*, CERN/LHCC **1997-33** (1997).

[59] CMS Collaboration: *"CMS Hadronic Calorimeter Project Technical Design Raport"*, CERN/LHCC **1997-31** (1997).

[60] CMS Collaboration: *"The CMS Muon Project Technical Design Raport"*, CERN/LHCC **1997-32** (1997).

[61] CMS Collaboration: *"CMS Trigger and Data-Acquisition System Project Technical Design Raport"*, CERN/LHCC **2000-38** (2000).

[62] CMS Collaboration: *"CMS Trigger and Data-Acquisition System Project Technical Design Raport Volume 2"*, CERN/LHCC **2002-26** (2002).

[63] J. Knobloch *et al.*: *"LHC Computing Grid: Technical Design Raport"*, CERN/LHCC **2005-24** (2005).

[64] CMS Collaboration: *"CMS Computing Technical Design Raport"*, CERN/LHCC **2005-23** (2005).

[65] S. Agostinelli *et al.*: *"GEANT4 - a Simulation Toolkit"*, Nucl. Instr. & Meth. A **506** (2003) 250–303.

[66] K. Holtmann: *"CMS Data Grid System Overview and Requirements"*, CMS Note **2001/037** (2001).

[67] V. Karimäki *et al.*: *"User's Guide for CMS Physics Generators Interface CMKIN Version 4"*, CMS IN **2004/052** (2004).

[68] *"CMS (reconstruction) software, http://cms-cpt-software.web.cern.ch/cms-cpt-software/General/"*, (30.10.2009) .

[69] T. Sjöstrand *et al.*: *"High-Energy-Physics Event Generation with PYTHIA 6.1"*, Comp. Phys. Comm. **135** (2001).

[70] T. Sjostrand, L. Lonnblad, S. Mrenna and P. Skands: *"Pythia 6.3 Physics and Manual"*, LU TP **03-38** (2003).

[71] H. Baer *et al.*: *"ISAJET 769: A Monte Carlo Event Generator for* pp, $\mathrm{p\bar{p}}$ *and* $\mathrm{e^+e^-}$ *Reactions"*, BNL-HET-03/028, UH-511-1038-03 (2004).

[72] G. Marchesini *et al.*: *"HERWIG - a Monte Carlo Event Generator For Simulating Hadron Emission Reactions With Interfering Gluons"*, Comp. Phys. Comm. **67** (1992) 465.

[73] F. Maltoni and T. Stelzer: *"MadEvent: Automatic Event Generation with MadGraph"*, JHEP-02 (2003).

[74] S. Jadach, Z. Was, R. Decker and J. Kuhn: *"The* τ *decay library TAUOLA: Version 2.4"*, Comp. Phys. Comm. **76** (1993) 361–380.

[75] M. Mangano *et al.*: *"ALPGEN, A Generator For Hard Multiparton Processes In Hadronic Collisions"*, CERN-TH **2002-129** (2002).

[76] S. Slabospitsky and L. Sonnenschein: *"TOPREX Generator (Version 3.25) Short Manual"*, Comp. Phys. Comm. **148** (2002) 87.

[77] A. Djouadi, J. Kalinowski and M. Spira: *"HDECAY: a Program for Higgs Boson Decays in the Standard Model and its Supersymmetric Extension"*, Comp. Phys. Commun. **108** (1998) 56.

[78] *"OSCAR: Object Oriented Simulation for CMS Analysis and Reconstruction, http://cmsdoc.cern.ch/oscar/"*, (30.10.2009) .

[79] V. Karimäki: *"CMSIM: CMS Simulation and Reconstruction Package, http://cmsdoc.cern.ch/cmsim/cmsim.html"*, (30.10.2009) .

[80] *"ORCA: Object-Oriented Reconstruction for CMS Analysis, http://cmsdoc.cern.ch/orca/"*, (30.10.2009) .

[81] V. Innocente, L. Silvestris and D. Stickland: *"CMS Software Architecture: Software Framework, Services and Percistency in High Level Trigger, Reconstruction and Analysis"*, Comp. Phys. Comm. **140** (2001) 31–44.

[82] CMS Collaboration: *"Plans for Jet Energy Corrections in CMS"*, CMS PAS **JME-07-002** (2007).

[83] S. Chekanov: *"Jet algorithms: a minireview"*, Proceedings of HCP02 (2002).

[84] G. Blazey *et al.*: *"Run II Jet Physics: Proceedings of the Run II QCD and Weak Boson Physics Workshop"*, FERMILAB-CONF-00-092-E (2000).

[85] S. Catani, Y. Dokshitzer, M. Seymour and B. Webber: *"Longitudinally Invariant Kt Clustering Algorithms for Hadron Hadron collisions"*, Nucl. Phys. B **406** (1993) 187–224.

[86] J. Butterworth, J. Couchman, B. Cox and B. Waugh: *"A C++ Implementation of the* $\mathrm{K_T}$ *clustering algorithm"*, Comp. Phys. Comm. **153** (2003) 85–96.

[87] S. Ellis and D. Soper: *"Successive combination jet algorithm for hadron collisions"*, Phys. Rev. D **48** (1993) 3160–3166.

[88] M. Cacciari and G. Salam: *"Dispelling the* N^3 *Myth for the* $\mathrm{K_T}$ *Jet-finder"*, Phys. Lett. B **641** (2006) 57–6.

[89] F. Blekman, A. Nikitenko and S. Wakefield: Private communication .

[90] P. Billoir: *"Progressive track recognition with a Kalman-like fitting procedure"*, Comp. Phys. Comm **57** (1989) 390–394.

[91] P. Billoir and S. Qian: *"Simultaneous pattern recognition and track fitting by the Kalman filtering method"*, Nucl. Instr. & Meth. A **294** (1990) 219–228.

[92] P. Billoir: *"Track fitting with multiple scattering: A new method"*, Nucl. Instr. & Meth. **225** (1984) 352–366.

[93] P. Billoir, R. Frühwirth and M. Regler: *"Track element merging strategy and vertex fitting in complex modular detectors"*, Nucl. Instr. & Meth. A **241** (1985) 115–131.

[94] R. Frühwirth: *"Application of Kalman filtering to track and vertex fitting"*, Nucl. Instr. & Meth. A **262** (1987) 444–450.

[95] S. Cucciarelli, M. Konecki, D. Kotlinski and T. Todorov: *"Track reconstruction, primary vertex finding and seed generation with the Pixel Detector"*, CMS Note **2006/026** (2006).

[96] S. Cucciarelli, M. Konecki, D. Kotlinski and T. Todorov: *"Track-Parameter Evaluation and Primary-Vertex Finding with the Pixel Detector"*, CMS Note **2003/026** (2003).

[97] P. Vanlaer, L. Barbone, N. D. Filippis, T. Speer, O. Buchmüller and F.-P. Schilling: *"Impact of CMS Silicon Tracker Misalignment on Track and Vertex Reconstruction"*, CMS Note **2006/029** (2006).

[98] R. Frühwirth *et al.*: *"Vertex reconstruction and track bundling at the LEP collider using robust algorithms"*, Comp. Phys. Commun. **96** (1996) 189–208.

[99] R. Frühwirth *et al.*: *"Adaptive Vertex Fitting"*, CMS Note **2007/008** (2007).

[100] T. Speer, K. Prokofiev, R. Fruehwirth, W. Waltenberger and P. Vanlaer: *"Vertex Fitting in the CMS Tracker"*, CMS Note **2006/032** (2006).

[101] A. Ostaptchouk *et al.*: *"The Alignment System of the CMS Tracker"*, CMS Note **2001/053** (2001).

[102] Belotelov *et al.*: *"Simulation of Misalignment Scenarios for CMS Tracking Devices"*, CMS Note **2006/008** (2006).

[103] R. Brun, F. Bruyant, M. Maire, A. McPherson and P. Zanarini: *"GEANT - a Detector Description and Simulation Tool"*, CERN Program Library Long Write-up **W5013** .

[104] S. Eno, W. Smith, S. Dasu, R. Kinnunen and A. Nikitenko: *"A Study of a First and Second Level Tau Trigger"*, CMS Note **2000/055** (2000).

[105] D. Kotlinski, A. Nikitenko and R. Kinnunen: *"Study of a Level-3 Tau Trigger with the Pixel Detector"*, CMS Note **2001/017** (2001).

[106] G.Bagliesi, S. Gennai and G.Squazzoni: *"A L2 Trigger for Tau Hadronic Decays with Tracker Isolation in the Low Luminosity Scenario"*, CMS Note **2002/018** (2002).

[107] S. Cucciarelli, D. Kotlinski and T. Todorov: *"Position Determination of Pixel Hits"*, CMS Note **2002/049** (2002).

[108] D. Denegri, R. Kinnunen and A. Nikitenko: *"Study of Calorimeter Calibration with Tau's in CMS"*, CMS Note **1997/039** (1997).

[109] S. Abdullin *et al.*: *"Summary of the CMS potential for the Higgs boson discovery"*, Eur.Phys.J. **C39S2** (2005) 41–61.

[110] A. Nikitenko, S. Kunori and R. Kinnunen: *"Missing Transverse Energy Measurement with Jet Energy Corrections"*, CMS Note **2001/040** (2001).

[111] R. Kinnunen and A. Nikitenko: *"Study of $\mathrm{H} \to \tau\tau$ with hadronic τ decays in CMS"*, CMS Note **2003/006** (2003).

[112] R. Kinnunen, S. Lehti, F. Moortgat, A. Nikitenko and M. Spira: *"Measurement of the H/A → tau tau Cross Section and possible Constraints on Tan(beta)"*, CMS Note **2004/027** (2004).

[113] CMS Collaboration: *"CMS Strategies for tau reconstruction and identification using particle flow techniques"*, CMS PAS **PFT-08-001** (2008).

[114] CMS Collaboration: *"Performance of Jet Algorithms in CMS"*, CMS PAS **JME-07-003** (2007).

[115] A. Rizzi, F. Palla and G. Segneri: *"Track Impact Parameter Based* b*-tagging Algorithm in CMS"*, CMS Note **2006/019** (2006).

[116] C. Weiser: *"A Combined Secondary Vertex Based* b*-tagging Algorithm in CMS"*, CMS Note **2006/014** (2006).

[117] A. Bocci, P. Demin, R. Ranieri and S. de Visscher: *"Tagging* b*- jets with electrons and muons in CMS"*, CMS Note **2006/043** (2006).

[118] S. Lowette, J. D'Hondt, J. Heyninck and P. Vanlaer: *"Offline Calibration of* b*-Jet Identification Efficiencies"*, CMS Note **2006/013** (2006).

[119] CMS Collaboration: *"Performance Measurement of b-tagging Algorithms Using Data containing Muons within Jets"*, CMS PAS **BTV-07-001** (2007).

[120] CMS Collaboration: *"*$\not\!\!E_{\mathrm{T}}$ *performance in CMS"*, CMS PAS **JME-07-001** (2007).

[121] S. Abdullin *et al.*: *"Summary of the CMS Potential for the Higgs Boson Discovery"*, CMS Note **2003/033** (2003).

[122] S. Gennai, A. Nikitenko and L. Wendland: *"Search for MSSM Heavy Neutral Higgs Boson in tau tau → two Jet Decay Mode"*, CMS Note **2006/126** (2006) 1–21.

[123] G. Degrassi, S. Heinemeyer, W. Hollik, P. Slavich and G. Weiglein: *"Towards high-precision predictions for the MSSM Higgs sector"*, Eur. Phys. J. C **28** (2003) 133–143.

[124] E. Boos *et al.*: *"CompHEP 4.4: Automatic computations from Lagrangians to events"*, Nucl. Instr. & Meth. A **76** (1993) 250–259.

[125] S. Gennai, F. Moortgat, L. Wendland, A. Nikitenko, S. Wakefield, G. Bagliesi, S. Dutta, A. Kalinowski, M. Konecki and D. Kotlinski: *"Tau jet reconstruction and tagging with CMS"*, Eur. Phys. J. C **46 s01** (2006) 1–21.

[126] A. Heister *et al.*: *"Jet reconstruction and performance in the CMS detector"*, Comp. Phys. Comm. **124** (2006) 76–89.

[127] M. Carena, S. Heinemeyer, C. Wagner and G. Weiglein: *"MSSM Higgs Boson Searches at the Tevatron and the LHC: Impact of Different Benchmark Scenarios"*, Eur. Phys. J. C **45** (2006) 797–814.

[128] H. Pi, P. Avery, D. Green, J. Rohlf and C. Tully: *"Measurement of Missing Transverse Energy with the CMS Detector at the LHC"*, CMS Note **2006/035** (2006).

[129] S. Bityukov, S. Erofeeva, N. Krasnikov and A. Nikitenko: *"Program for evaluation of significance, confidence intervals and limits by direct calculation of probabilities"*, Proceedings of PhyStat 2005 (2005).

[130] R. Kinnunen: *"Search for the heavy charged MSSM Higgs bosons with the* $\mathrm{H}^{\pm} \to \tau^{\pm}\nu_{\tau}$ *decay mode in fully hadronic final state, revised version"*, CMS Note **2006/100** (2006).

[131] E. L. Berger, T. Han, J. Jiang and T. Plehn: *"Associated production of a top quark and a charged Higgs boson"*, Phys. Rev. D **71** (2005).

[132] V. Konoplianikov, O. Kodolova and A. Ulyanov: *"Jet Calibration Using γ+jet Events in the CMS Detector"*, CMS Note **2006/042** (2006).

[133] G. Davatz, M. Dittmar and A.-S. Giolo-Nicollerat: *"Standard Model Higgs Discovery Potential of CMS in $H \to WW \to l\nu l\nu$ channel"*, CMS Note **2006/047** (2006).

[134] A. Kalinowski and A. Nikitenko: *"Measurement of the τ tag efficiency using the $\mathrm{Z} \to \tau\tau \to \mu + \mathrm{hadrons} + \mathrm{X}$ events"*, CMS Note **2006/074** (2006).

[135] D. Chakraborty, J. Konigsberg and D. Rainwater: *"Top-quark Physics"*, Ann. Rev. Nucl. Part. Sci. **53** (2003) 301–351.

[136] M. Baarmand, M. Hashemi and A. Nikitenko: *"Light Charged Higgs Discovey Potential of CMS in the $\mathrm{H}^{\pm} \to \tau\nu$ Decay with Single Lepton Trigger"*, CMS Note **2006/056** (2006).

[137] M. Hashemi, S. Heinemeyer, R. Kinnunen, A. Nikitenko and G. Weiglein: *"Charged Higgs Bosons in the MSSM at CMS: Discovery Reach and Parameter Dependence"*, hep-ph/08041228 (2008).

[138] H. Lai *et al.*: *"Global QCD Analysis of Parton Structure of the Nucleon: CTEQ5 Parton Distributions"*, Eur. Phys. J. C **12** (2000) 375–392.

[139] T. Plehn: *"Charged Higgs boson production in bottom gluon fusion"*, Phys. Rev. D **67** (2003).

[140] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau and A. D. Polosa: *"ALPGEN, a generator for hard multiparton processes in hadronic collisions"*, JHEP 0307 **001** (2003).

[141] V. Bartsch and G. Quast: Private communication .